



Etude par modélisation de dynamique moléculaire et spectroscopie RMN des déformations induites par la coordination du cisplatine sur l'ADN

Stéphane Téletchéa

► To cite this version:

Stéphane Téletchéa. Etude par modélisation de dynamique moléculaire et spectroscopie RMN des déformations induites par la coordination du cisplatine sur l'ADN. Médicaments. Université René Descartes - Paris V, 2005. Français. NNT: . tel-00593727

HAL Id: tel-00593727

<https://theses.hal.science/tel-00593727>

Submitted on 17 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris V - René Descartes
Ecole doctorale du médicament

Thèse de doctorat
pour obtenir le grade de :
DOCTEUR DE L'UNIVERSITE PARIS 5

Présentée par :

Stéphane Téletchéa

**« Etude par modélisation de
dynamique moléculaire et
spectroscopie RMN des déformations
induites par la coordination du
cisplatine sur l'ADN »**

Soutenue le 27 septembre 2005 devant le jury composé de :

Rapporteur	: Pr. F. Vovelle
Rapporteur	: Dr. S. Femandjian
Examineur	: Pr. J-S. Sun
Examineur	: Dr. N. Gresh
Directeur de thèse	: Dr. J. Kozelka

A mes parents

Remerciements

Je tiens à remercier J. Kozelka pour m'avoir permis d'obtenir ce diplôme de doctorat. Il est le fruit d'un travail enrichi de collaborations, de nombreuses expériences humaines et scientifiques, en France comme à l'international. J'ai pu, à travers un projet ambitieux, aborder de nombreuses techniques de modélisation et les valider par une approche expérimentale.

Je remercie les membres du jury pour avoir accepté de juger mon travail de thèse et de m'avoir accordé leur plus grande confiance pour mon avenir scientifique.

Je remercie ceux qui m'ont accueilli au sein des locaux des Saints-Pères, MM. Chotard et Mansuy pour l'unité UMR8601, M. Jore, Mme Garbay, Mmes Curmi et Goddé pour l'Université Paris V. Ils ont facilité les démarches administratives (changement d'école doctorale, financement étranger, ...) et c'est grâce à leur contribution que j'ai pu arriver au sein du laboratoire et de l'école doctorale du médicament dans de bonnes conditions.

Je remercie tous les membres de l'UMR 8601 pour les bons moments qu'ils m'ont permis de vivre, tout particulièrement Sophie Bombard et Diana Over. Grâce à leur enseignement, j'ai pu me rappeler pourquoi j'avais décidé d'abandonner lâchement le travail expérimental en biologie (tout du moins réduire sa part dans mon travail de recherche à une fraction infime) mais aussi découvrir de nombreux domaines mal connus ou inconnus, biologiques et chimiques. A travers l'enseignement de Diana et de ses collaborateurs (que je remercie ici encore une fois), j'ai pu mettre à jour mes connaissances en chimie organique et découvrir le monde merveilleux de la chimie inorganique, bien plus riche que je ne l'aurai imaginé.

Je remercie aussi les étudiants et post-doctorants qui ont partagé la même aventure que moi durant ces longues années : Olivier Delalande, Markus Drumm, Véronique Montjardet-Bas, Sandrine Perrier, Miguel-Angel Elizondo-Riojas. A travers nos longues discussions sur nos thèmes de prédilection (la thèse, la supervision, parfois la maladie ...) de nombreux moments ont été partagés, très riches en enseignement, qui m'ont

permis de mûrir mes perspectives de vie. Nous avons maintenant tous suivi des voies parallèles et sans nul doute nous serons amenés à nous rencontrer dans notre futur professionnel.

Un grand merci tout particulier à Miguel-Angel Elizondo-Riojas pour m'avoir fait bénéficier de sa grande expérience dans le domaine de la modélisation moléculaire et des subtilités des différents programmes d'AMBER. Un autre grand merci à Véronique Montjardet-Bas pour avoir passé de longues heures à m'enseigner le fonctionnement de nos chères vieilles HPLC et surtout les subtilités de purification et d'analyse des adduits platinés.

Dans le cadre des collaborations qui ont fait la richesse de ce travail de thèse, un grand merci à T. Skauge pour son aide précieuse dans l'analyse RMN et à E. Sletten pour l'accueil au sein de son laboratoire durant mon séjour Norvégien. *Takk for det glimrende arbeidet vi har gjort sammen!*

Merci à B. Hartmann et D. Djuranović pour la collaboration sur la description fine du comportement dynamique de l'ADN et sur la caractérisation des sous-états BI et BII.

Merci aux membres du laboratoire de Leiden (Pays-Bas), J. Reedijk, S. Komeda, J-M. Teuben.

Cette thèse n'aurait probablement pas pu être réalisée sans le financement apporté par « *The Association for International Cancer Research (AICR)* », j'espère avoir apporté ma modeste contribution à la lutte contre le cancer. Je pense avoir utilisé honnêtement et sérieusement l'argent des personnes qui font confiance à cette association pour lutter contre le cancer.

Le financement indirect de la fin de ma thèse a été l'occasion de réaliser un rêve, travailler chez un acteur majeur des Logiciels Libres : Mandrakesoft devenu Mandriva durant ma période de travail. Cette perspective saugrenue en début de thèse a été l'occasion de valider ma profonde conviction que les Logiciels Libres et que les acteurs qui y contribuent, qu'ils fussent simples particuliers ou des sociétés cotées en bourse, ont un rôle majeur à jouer dans l'évolution de l'Humanité. J'ai choisi depuis d'orienter ma carrière vers la recherche mais je compte bien continuer à m'impliquer dans cette formidable communauté. Merci à tous ceux que j'ai rencontrés durant ces nombreuses expériences, qui m'ont permis d'avancer quand je bloquais sur un script, un programme, une permission de fichier ...

Le Conseil général de Saint-Pierre et Miquelon et indirectement tous les habitants de l'archipel qui m'ont permis d'arriver au terme de ces études à travers une politique volontariste affichée et ambitieuse. Bien qu'étant promis à un avenir hors de nos magnifiques îles, je suis très fier de venir d'un pays de marins. Nos ancêtres avaient fait le choix du déracinement vers l'archipel pour 'tenter l'aventure', je me retrouve bien malgré moi contraint de faire le chemin inverse. J'espère que ce travail de thèse permettra par sa modeste contribution de faire mieux connaître l'archipel, en tous cas, moi, je ne perds jamais une occasion de faire la publicité de la beauté de celui-ci et de la chaleur humaine de ses habitants.

Les derniers remerciements iront à mes parents et à ma famille pour leur soutien indéfectible, et à Christelle qui a patiemment passé de longues heures à me supporter et à m'aimer.

Résumé

Le cisplatine (ou cis-diammine, dichloro-platine) est l'un des composés chimiques les plus utilisés actuellement en chimiothérapie anticancéreuse. Depuis la description de ses propriétés anticancéreuses par B. Rosenberg en 1965 de nombreux travaux ont été effectués afin de décrire le mécanisme d'action lui conférant ses propriétés antitumorales.

A travers une approche originale couplant modélisation et travail expérimental, les recherches réalisées durant ma thèse ont permis d'élucider le comportement dynamique d'un adduit platiné sur la séquence 5'-GCCG*G*GTCGC-3' / 5'-GCGACCCGGC-3' (G* représente une guanine platinée). Cette structure a été comparée à celle de l'adduit ADN-cisplatine déterminée précédemment au laboratoire sur la séquence G*G*A. Nous avons ainsi étudié l'influence d'une guanine adjacente en 3' au pontage GG-Pt sur la structure de l'adduit. Il s'agit de la première étude structurale sur un adduit du cisplatine avec la séquence GGG. Même si l'affinité de la séquence GGG et des sites contenant Gn ($n \geq 3$) pour le platine (II) est connue depuis longtemps, ses adduits avec le cisplatine n'ont pas encore été étudiés par RMN à cause des problèmes posés par leur purification.

Cette étude de l'adduit G*G*G par RMN a été confrontée à la description dynamique de cet adduit, calculée par simulation. La paramétrisation du champ de force parm 98 a été spécifiquement affinée pour mieux décrire l'environnement de l'atome du platine. La confrontation entre la simulation de l'adduit G*G*G-Pt et les données issues de l'étude par RMN a permis de valider notre paramétrisation.

Pour déterminer les proportions précises des sous-états BI et BII de l'ADN, une méthode novatrice a été mise au point. Celle-ci est basée sur la combinaison de quatre distances inter-protons H2''(n)-H8(n), H1'(n)-H6/8(n+1), H2'(n)-H6/8(n+1) et H2''(n)-H6/8(n+1) qui permet de discriminer les deux sous-conformations.

Ces améliorations du champ de force et de la méthode de détermination des sous-conformations BI/BII ont permis la description fine du comportement de l'adduit couplé à l'ADN, ce qui nous a servi pour étudier le mécanisme anti-tumoral du cisplatine. En effet la reconnaissance de l'ensemble cisplatine-ADN par une protéine (Lymphoïd Enhanced Factor I - LEF I) pourrait activer les voies métaboliques de la cellule cancé-

reuse conduisant vers l'apoptose ou vers la réparation de la tumeur. La simulation de l'ensemble ADN-protéine (sans cisplatine) a permis de présenter le mode de reconnaissance de la protéine sur la déformation ainsi que la mise en évidence de l'implication d'une molécule d'eau dans celui-ci.

Les études sur le cisplatine fixé sur son ADN cible nous ayant apporté de nombreuses connaissances sur les déformations engendrées, nous avons validé la déformation structurale formée par un autre complexe de platine, le pyrazolato-bis-platine. Ce composé a été conçu de novo pour induire une déformation faible de son ADN cible afin de provoquer une réponse cellulaire différente de celle engendrée par le cisplatine. La simulation a indiqué que ce complexe induit une faible courbure de l'ADN et une déformation globale différente de celle du cisplatine, ce qui exclut le même mode de reconnaissance. Comme expérimentalement le complexe pyrazolato possède une activité anti-tumorale, les simulations effectuées suggèrent donc que le mode d'action au niveau cellulaire est différent du cisplatine.

Les travaux réalisés lors de ma thèse ont permis d'améliorer la compréhension des déformations ADN-cisplatine, ADN-cisplatine-protéine et de perfectionner la description des composés platinés dans le champ de force parm 98 (puis parm 99) du logiciel de modélisation moléculaire AMBER.

Abstract

Cisplatin (cis-diammin dichloro platin) is one of the most widely used anti cancer chemotherapy drug. Its anticancer activity has been first described in 1965 by B. Rosenberg. Since then, many studies have been undertaken to decipher its mechanism of action.

From an original approach linking molecular dynamics simulation and experimental studies, my PhD allowed a dynamic description of a platinated adduct on the sequence 5'-GCCG*G*GTCGC-3' / 5'-GCGACCCGGC-3' (where G* is a platinated guanine). The resulting structure has been studied previously in our laboratory on the identical sequence but with an A-T base pair on the 3' side of the platinated guanines (subsequently named G*G*A). Therefore we studied the structural implications of an A to G substitution in the 3' region on the adduct structure. This is the first structural study on a cisplatin adduct containing three adjacent guanines. GGG sequences and Gn ($n \geq 3$) are known to be favored targets for platin (II) but their adducts have not been studied by NMR since their purification are problematic.

This G*G*G NMR study has been compared to its dynamics simulation, where the parm 98 force field has been specifically optimized for taking into account the platin coordination sphere. The parametrisation has been cross-validated by the NMR experiment.

To calculate the proportion of BI/BII DNA substates, a novel method has been set up. Four distances, namely H2''(n)-H8(n), H1'(n)-H6/8(n+1), H2'(n)-H6/8(n+1) and H2''(n)-H6/8(n+1) allow to discriminate between the two substates and to calculate the proportion of each substate.

The force field optimizations and the BI/BII DNA substates discrimination method allowed a precise description of the cisplatin-DNA adduct, allowing to study one possible way of anti-cancer efficiency. The complex cisplatin-DNA-Protein (notably LEF I - Lymphoid Enhanced Factor I) is a plausible candidate to explain why the cancerous cell elects apoptosis or tumor repair. The protein-DNA (without cisplatin) ensemble simulation has allowed to describe the recognition path of the DNA deformation by the protein and the implication of a water molecule in the recognition mechanism.

All these studies on cisplatin-DNA allowed us to get enough knowledge on the

resulting deformations. Therefore, we then studied the structural deformation induced by another platin-related complex, the pyrazolate-bis-platine. This complex has been designed to obtain a small DNA deformation, in order to get a different response than the one of cisplatin. The molecular dynamics simulation has indicated it induces a smaller kink and a different global deformation than cisplatin thus proving a different recognition mode. Since it is known by *in vitro* studies that the pyrazolate-bis-platine owns an anti-tumoral activity, the dynamics simulations suggest its cellular mechanism is different than the one of cisplatin.

These PhD studies have allowed a better understanding of DNA-cisplatin, DNA-cisplatin-protein dynamic description and to improve the definition of platinated complex for force fields parm 98 (and parm 99) of the molecular dynamics simulation suite AMBER.

Table des matières

Introduction générale	1
1 Les complexes de platine	3
1.1 Le cisplatine	3
1.2 Autres complexes actifs contenant du platine	4
1.3 Action supposée du cisplatine sur les cellules cancéreuses	7
1.3.1 Adduits formés	7
1.3.2 Reconnaissance des adduits du platine	7
1.3.3 Inhibition de la réplication	9
1.3.4 Inhibition de la transcription	9
1.3.5 Réparation des adduits du cisplatine	10
1.3.6 Propriétés mutagènes des adduits du cisplatine	10
1.3.7 Résistance au cisplatine	11
2 L'ADN	13
2.1 Les bases	13
2.2 Le squelette phosphodiester	16
2.3 Le sucre furanose	17
2.4 Les paramètres hélicoïdaux	19
2.5 Polymorphisme de l'ADN	22
2.5.1 Forme B	23
2.5.2 Forme A	24
2.5.3 Forme Z	26
2.6 Conclusion	28
3 Simulation de dynamique moléculaire	29
3.1 Méthodologie de la Dynamique Moléculaire	29
3.2 Simulation en solvant explicite	31
3.2.1 Conditions périodiques et seuil de calcul	32
3.2.2 Restrictions sur la longueur des liaisons	33
3.3 Le champ de force d'AMBER	34

3.4	Mise en place d'une simulation de dynamique moléculaire	35
3.4.1	Préparation du composé	35
3.4.2	Chauffage	37
3.4.3	Equilibration	37
3.4.4	Production	37
3.4.5	Analyse	37
4	Objectifs du travail	39
I	Détection de la conformation BII	43
II	RMN de l'adduit platiné G*G*G	69
1	Préparation de l'adduit platiné	73
2	RMN de l'adduit	75
2.1	Analyse de l'échantillon	75
2.2	Déplacements chimiques observés	76
2.3	Différences entre platiné et non platiné	79
2.4	Attribution des sucres	80
2.4.1	Phase des sucres	80
2.4.2	Informations issues des constantes de couplages $J_{1'2'}$ et $J_{1'2''}$. .	81
2.4.3	Indications sur les sucres à partir des expériences NOESY . . .	82
2.5	Distances RMN	84
2.6	Réarrangement de l'adduit	84
3	Dynamique moléculaire	87
3.1	Préparation de la dynamique	88
3.2	Analyse de la dynamique	88
3.2.1	Caractéristiques générales issues de la simulation	89
3.2.2	Conformation des sucres	89
3.2.3	Conformation des phosphates	90
3.2.4	Comparaison des distances RMN avec le modèle dynamique . .	90
3.2.5	Validation du modèle simulé par les données expérimentales . .	90
4	Conclusion	93

III	Reconnaissance plausible d'un ADN platiné par la LEF-I	95
IV	Etude du complexe pyrazolato-bisplatine	135
V	Les composés platinés dans AMBER	197
1	Caractéristiques chimiques du cisplatine	201
2	Adaptation du champ de force parm 98	203
3	Déformations induites sur le squelette	207
3.1	Présentation des systèmes d'étude	208
3.2	Le sucre platiné	209
3.3	L'angle ϵ	212
4	Intégration dans le champ de force	221
4.1	Préparation du composé	221
4.2	Simulation de dynamique moléculaire	222
4.3	Analyse de la dynamique	222
5	Conclusion	227
VI	Conclusion générale	229
	Bibliographie	237
	Annexes	259

Table des figures

1.1	Le cisplatine	3
1.2	Le carboplatine [Pt(NH ₃) ₂ (cyclobutane-1,1-dicarboxylate)].	5
1.3	Les complexes de « Hollis »	6
1.4	Les complexes de « Farrell »	6
2.1	Les bases nucléotidiques.	14
2.2	Appariements Watson-Crick et définition des sillons	15
2.3	Appariement selon Hoogsteen entre l'adénine et la thymine, dans la structure cristalline formée par la paire 9-méthyladénine/1-méthylthymine [194]. . . .	15
2.4	Le squelette phosphodiester et les angles le définissant.	16
2.5	Le sucre furanose et les angles qui le décrivent.	17
2.6	Cercle de pseudo-rotation. Les signes des cinq angles dièdres endo-cycliques sont indiqués à l'intérieur du cycle furanose [3]	19
2.7	Représentation schématique des plissements C2'-endo (gauche) et C3'-endo (droite)	19
2.8	Configuration syn et anti de l'angle glycosidique χ . Le sucre et la base sont schématisés en bleu et en vert, respectivement.	20
2.9	Paramètres hélicoïdaux	21
2.10	Paramètres intra-paires de bases	21
2.11	Paramètres inter-paires de bases	22
2.12	ADN B. P : petit sillon, G : grand sillon.	25
2.13	ADN A	26
2.14	ADN Z	27
4.1	Représentation schématique du décanucléotide d(GCCG*G*ATCGC)·d(GCG ATCCGGC) avec l'adduit GG- <i>cis</i> -Pt(NH ₃) ₂ sur les guanines G* analysé par spectroscopie RMN par Herman et ses collaborateurs [87].	40

- 2 projection bi-dimensionnelle de la différence entre les distances H2'' (n)-H6/H8 (n) et H2'' (n)-H6/H8 (n+1) contre la distance H1' (n)-H6/H8 (n+1). Chaque point représente une des 26978 conformations dinucléotiques issues des simulations de dynamique moléculaire. Les couleurs indiquent les sous-conformations des pas. En bleu : BI, sucre 5' S, phosphate 3' BI. En Orange : BI, sucre 5' E, phosphate 3' BI. En noir : BI, sucre 5' N, phosphate 3' BI. En magenta : BI, phosphate 3' BII. En rouge : BII, phosphate du brin complémentaire en BI. En vert : BII, phosphate du brin complémentaire en BII. Pour déterminer le plissement du sucre, les critères suivants ont été appliqués pour l'angle de la phase P : N, $P < 60^\circ$; E, $60^\circ \leq P \leq 115^\circ$; S, $115 \leq P \leq 220^\circ$. Le signe '+' indique le barycentre de chaque conformation, X_{BI} , X_{BII} la projection de ceux-ci. Le point P présente à titre d'exemple la projection obtenue pour le pas P donné. 47
- 1.1 Evolution de la réaction du décanucléotide GCCGGGTCGC avec le cisplatine (cis-[PtCl₂(¹⁵NH₃)₂]). 1 équivalent du cisplatine a été ajouté au temps t=0, puis addition de 1 équivalent à 20h puis 0,5 équivalent à 36h et 43h. Le gradient HPLC utilisé commence à 2% d'acétonitrile puis augmente de manière exponentielle à 20 % jusqu'à 25 minutes (programme 3 du logiciel Waters Millenium32 ver. 3.05.01). Les chélates sont séparés en HPLC au bout de 10 minutes, soit un pourcentage d'acétonitrile de 8%. Le passage de 8 à 20 % permet de rincer la colonne pour effectuer une nouvelle analyse. Le solvant organique est constitué d'acétonitrile (mais est dilué dans 50% d'eau pour faciliter l'analyse), le solvant aqueux contient l'acétate d'ammonium 0,5 M ajusté à pH 4,5. 74
- 2.1 Vue étendue du spectre NOESY de la région G*G*G, d[(GCCG*G*GTCGC) · (GCGACCCGGC)]. La vue présente le chemin d'attribution du brin platiné dans la région aromatique-anomérique. Les cercles en pointillés indiquent les pics manquants. Les pics qui ne font pas partie du chemin d'attribution appartiennent au brin complémentaire et au duplex GG*G*. La concentration de l'oligonucléotide est de 0,50 mM dans 1 M NaClO₄, 30 mM tampon phosphate, pH 6,0 en D₂O à 305 K. 76

2.2	Vue étendue du spectre NOESY de la région G*G*G, d[(GCCG*G*GTCGC) · (GCGACCCGGC)]. La vue présente le chemin d'attribution du brin complémentaire non platiné dans la région aromatique-anomérique. Les pics qui ne font pas partie du chemin d'attribution appartiennent au brin platiné et au duplex GG*G*. La concentration de l'oligonucléotide est de 0,50 mM dans 1 M NaClO ₄ , 30 mM tampon phosphate, pH 6,0 en D ₂ O à 305 K.	77
2.3	Différence entre les déplacements chimiques pour les protons non échangeables entre G*G*G et G*G*A. Les spectres ont été enregistrés à 305K (G*G*G) et 298K (G*G*A).	80
2.4	Différence entre les déplacements chimiques pour les protons non échangeables entre G*G*G et GGG. Les spectres ont été enregistrés à 305K (G*G*G).	80
2.5	Déplacements chimiques du phosphore où la présence de G*G*G et GG*G* est détectée. (a) spectre enregistré 100 jours après la platination, les deux adduits sont présents. (b) spectre de l'oligonucléotide non platiné. Les déplacements chimiques sont présentés avec pour référence l'acide phosphorique à 85 %, l'enregistrement a été effectué à 305 K. . .	85
2.6	diagramme cinétique du réarrangement de G*G*G vers GG*G*. Jours* : le temps présenté est t_{eff} , le calcul de ce temps est décrit dans le texte.	86
3.1	Représentation <i>NewRibbon</i> et <i>Licorice</i> [95] de la structure moyennée et minimisée issue de la simulation de dynamique moléculaire. Les guanines platinées sont en bleu, les ammines du cisplatine en jaune, le platine en cyan.	92
1.1	Le cisplatine	201
2.1	Modèle de représentation du cisplatine dans le champ de force d'AMBER.	204
3.1	Modification apportée pour singulariser le sucre de la guanine platinée. Le type d'atome CT devient CP pour les atomes du cycle.	210
3.2	Influence de la valeur de la constante pour l'angle de torsion O-C-C-O sur la phase de 5G (rouge), la phase de 3G (vert) et la distance ${}^5G(H_1) - {}^3G(H_1)$ (noir). Les tracés sont moyennés sur 5 ps pour plus de lisibilité.	211
3.3	Modification apportée pour influencer uniquement sur l'angle ϵ . Le carbone C ₂ ' est maintenant de type CS.	213
3.4	Evolution du potentiel d'énergie de torsion autour de l'angle CS-CP-OS-P en fonction de la constante PK_{CCOP} d'énergie utilisée.	214

3.5	Prolongation de la simulation de dynamique moléculaire avec parm 98. La phase des sucres est représentée en vert, l'angle de torsion ϵ est en noir, l'angle de torsion ζ est en rouge. Les 200 premières ps délimitées par la ligne verticale bleue présentent le comportement dynamique avec parm 98 natif. Les tracés sont moyennés sur 5 ps pour plus de lisibilité.	216
3.6	Prolongation de la simulation de dynamique moléculaire avec parm 98 puis modification des sucres platinés et du potentiel de torsion avec $PK_{CCOP} = 1,00 \text{ kcal.mol}^{-1}$ pour ϵ . La phase des sucres est représentée en vert, l'angle de torsion ϵ est en noir, l'angle de torsion ζ est en rouge. Les 200 premières ps délimitées par la ligne verticale bleue présentent le comportement dynamique avec parm 98 natif. Les tracés sont moyennés sur 5 ps pour plus de lisibilité.	217
3.7	Prolongation de la simulation de dynamique moléculaire avec parm 98 puis modification des sucres platinés et du potentiel de torsion avec $PK_{CCOP} = 0,50 \text{ kcal.mol}^{-1}$ pour ϵ . La phase des sucres est représentée en vert, l'angle de torsion ϵ est en noir, l'angle de torsion ζ est en rouge. Les 200 premières ps délimitées par la ligne verticale bleue présentent le comportement dynamique avec parm 98 natif. Les tracés sont moyennés sur 5 ps pour plus de lisibilité.	218
3.8	Prolongation de la simulation de dynamique moléculaire sous parm 98 avec le potentiel de torsion modifié $PK_{CCOP} = 1,00 \text{ kcal.mol}^{-1}$ pour ϵ mais sans modification des sucres platinés. La phase des sucres est représentée en vert, l'angle de torsion ϵ est en noir, l'angle de torsion ζ est en rouge. Les 200 premières ps délimitées par la ligne verticale bleue présentent le comportement dynamique avec parm 98 natif. Les tracés sont moyennés sur 5 ps pour plus de lisibilité.	219
4.1	Contrôle du bon fonctionnement de la dynamique, en utilisant le programme standard <code>process_mdout.perl</code> fourni dans AMBER.	224
4.2	Paramètres principaux issus de la simulation de dynamique molécule de GCCG* G*GTCGC. Les noms d'atomes sont ceux de la figure 2.1, page 204. C3 et G6 sont respectivement les bases situées en 5' et en 3' de l'adduit platiné. C15 est la cytosine complémentaire de $3'G$, C16 est la cytosine complémentaire de $5'G$. Les comparaisons avec les données expérimentales sont présentées dans le chapitre II. Les tracés sont moyennés sur 10 ps pour plus de clarté.	225

Liste des tableaux

1.2	Adduits majeurs provoqués par le cisplatine et caractéristiques principales (EG : électrophorèse sur gel, RMN : résonance magnétique nucléaire, RX : diffraction aux rayons X)	7
2.1	Paramètres structuraux des trois familles de l'ADN. Les translations sont en Å et les rotations en degrés. Pour l'ADN Z, les valeurs pour les paires YpR puis pour les paires RpY sont données. Elles reflètent la symétrie dinucléotidique de cette forme [81].	24
3.1	Déroulement du calcul d'un pas d'intégration de dynamique moléculaire.	31
1	Récapitulatif des 21 distances mesurées sur les structures issues des données cristallographiques et des simulations de dynamique moléculaire. Les distances indiquées en rouge sont celles présentant le plus d'intérêt dans le cadre de notre étude.	46
2.1	Déplacements chimiques pour ^1H du duplex platiné $\text{G}^*\text{G}^*\text{G}$. Tous les spectres ont été enregistrés à 305 K à l'exception des spectres des protons échangeables pour $\text{G}^*\text{G}^*\text{G}$ enregistrés à 281 K. Les déplacements ont pour référence le TSP (Trisilyl Méthyl Phosphate).	78
2.2	Déplacements chimiques pour ^1H du duplex non-platiné GGG. Tous les spectres ont été enregistrés à 305 K. Les déplacements ont pour référence le TSP (Trisilyl Méthyl Phosphate).	79

2.3	Couplages spin-spin $J_{1'2'}$, $J_{1'2''}$ et $\Sigma_{1'}$ déterminés à partir des spectres DQF-COSY, distances inter-protons H2'-H6/8 et H3'-H6/8 déterminées à partir des spectres NOESY, et ratio $\frac{H2'-H6/8}{H3'-H6/8}$ pour le duplex platiné G*G*G. Une estimation de la conformation du sucre, calculée à partir du ratio $\frac{H2'-H6/8}{H3'-H6/8}$, est présentée dans la dernière colonne (voir le texte pour les explications du calcul). ^a Les distances inter protons H2'-H6/8 et H3'-H6/8 ont été moyennées à partir des trois spectres NOESY : temps de mélange de 200 ms à 305 K en D ₂ O, temps de mélange de 80 ms à 305K en D ₂ O et temps de mélange de 200 ms à 281 K en H ₂ O. Toutes les distances n'ont pas pu être mesurées pour chaque temps de mélange.	82
2.4	Déplacements chimiques pour ¹ H du duplex GG*G*. Tous les spectres ont été enregistrés à 305 K à l'exception des spectres pour les protons échangeables qui ont été enregistrés à 281 K. Les déplacements ont pour référence le TSP (Trisilyl Méthyl Phosphate).	84
3.1	Comparaison entre les conformations des sucres observées en RMN et celles calculées à partir de la simulation de dynamique moléculaire.	89
2.1	Caractéristiques majeures de la paramétrisation du cisplatine dans le champ de force d'AMBER. HT : atome d'hydrogène relié aux ammines platinées. Le détail de chaque résidu est présenté en annexe, page 257.	204
4.1	Protocole de simulation de dynamique moléculaire.	223

Introduction générale

Chapitre 1

Les complexes de platine

1.1 Le cisplatine

Le cisplatine (ou cis-diammine, dichloro-platine) est l'un des composés chimiques les plus utilisés actuellement en chimiothérapie anticancéreuse (figure 1.1).

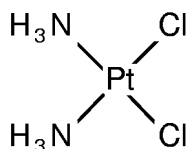


FIG. 1.1 – Le cisplatine

La première description de sa synthèse date de 1845 [147]. En 1965, Barnett Rosenberg, en étudiant l'effet du courant électrique sur la croissance d'*Escherichia Coli*, observe que les bactéries ne se divisent plus mais deviennent filamenteuses lorsqu'elles sont placées dans un champ électrique fourni par des électrodes de platine [162]. Les espèces chimiques responsables de cette croissance anormale sont identifiées comme des sels de platine, rhodium et ruthénium. L'analyse des espèces en solution précise ensuite que ce sont majoritairement les complexes *cis*-[PtCl₄(NH₃)₂] et *cis*-[PtCl₂(NH₃)₂], formés par la lente dissolution par électrolyse des électrodes de platine contenant du chlorure d'ammonium, qui sont actifs [161]. Poursuivant l'analyse de ce comportement anormal, il en a surtout déduit de cette observation que si ces composés étaient capables d'interrompre la division cellulaire sans tuer les bactéries, ils pourraient être susceptibles d'arrêter la prolifération cellulaire des cellules cancéreuses chez les mammifères et l'homme en particulier. Cette hypothèse a été validée ensuite, le complexe *cis*-[PtCl₂(NH₃)₂] se révélant être le plus actif des deux sur les cellules testées [163].

Les premiers essais cliniques commencèrent en 1971 et le cisplatine est utilisé depuis 1979 dans les thérapies anti-cancéreuses. Son action est la plus efficace dans le cas des

cancers des testicules et des ovaires [120] et dans une moindre mesure pour les cancers du nasopharynx, de la vessie, des os et de certains cancers des poumons [10, 152].

Cependant le cisplatine se révèle être très toxique, ce qui a dans un premier temps empêché son approbation par la Food and Drug Administration. Les effets secondaires principaux sont de deux types :

- les nausées et vomissements qui apparaissent une heure après l'administration du médicament et durent quelques heures ;
- les toxicités à long terme, dont la néphrotoxicité due à une accumulation du cisplatine dans les reins, la neurotoxicité, la myélotoxicité et l'ototoxicité (perte de capacité auditive) [152].

Enfin une résistance au traitement par le cisplatine peut apparaître, comme cela se produit en général dans les chimiothérapies.

Ces effets secondaires sont maintenant considérablement diminués par l'hydratation préalable du patient (1 à 2 litres de sérum physiologique), l'infusion du cisplatine dilué dans des solutions salines (0,3-0,5 M NaCl) contenant 5 % de dextrose [152]. Pour contrer la néphrotoxicité, des composés soufrés sont utilisés (comme le thiosulfate de sodium, le diéthylcarbamate) pour réduire la fixation covalente du cisplatine sur les protéines rénales.

1.2 Autres complexes actifs contenant du platine

Lors de la recherche de nouveaux complexes de platine présentant une meilleure efficacité thérapeutique, il a été remarqué que les composés actifs avaient certaines caractéristiques communes [42, 41], ce qui a permis d'établir une première relation structure-activité :

- une géométrie « plan-carré » pour les complexes de Pt(II) ou octaédrique pour les complexes du Pt(IV) ;
- la présence de deux ligands inertes dans les conditions physiologiques (en général amines) en position *cis* ;
- la présence de deux ligands monodentates en position *cis* ou d'un ligand bidentate, labile ;
- une labilité intermédiaire des groupes partants :

1. si les ligands sont trop labiles (comme NO_3^- ou SO_4^{2-}), le complexe présente une forte toxicité car il réagit avec les protéines du plasma et les protéines rénales [11, 208],
2. si les ligands sont trop peu labiles (comme I^- ou SCN^-), le complexe a une faible efficacité [42, 125].

- la présence d'au moins un atome d'hydrogène sur les groupes non partants [42, 87, 158].

En tenant compte de ces caractéristiques, des complexes dits de deuxième génération ont été développés dans les années 70-80, notamment pour limiter la toxicité rénale due à l'aquation prématurée des ligands chloro. Parmi les nombreux composés synthétisés, plus de 15 sont entrés en phase d'essai clinique mais un seul est administré actuellement en routine : le carboplatine (figure 1.2).

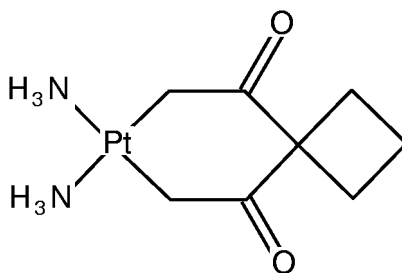


FIG. 1.2 – Le carboplatine $[\text{Pt}(\text{NH}_3)_2 (\text{cyclobutane-1,1-dicarboxylate})]$.

Le spectre d'activité du carboplatine est identique à celui du cisplatine. Sa toxicité est inférieure (son ligand dicarboxylatocyclobutane est moins labile que les ligands chlorure du cisplatine) mais son activité est aussi moindre. Pour obtenir une efficacité équivalente, il faut administrer des doses dix fois supérieures à celles du *cis*-DDP [134]. De plus, il présente une résistance croisée avec le cisplatine [2].

Depuis quelques années, de nombreux complexes de platine ne remplissant pas toutes les caractéristiques présentées précédemment, ont révélé une activité *in vitro* et *in vivo* :

- les complexes dits « de Hollis ». Ils comportent trois groupements inertes - deux ligands ammine en position *cis* et un ligand amine aromatique ou pyrimidine - et possèdent une activité *in vivo*, contrairement aux complexes $[\text{PtCl}(\text{NH}_3)_3]$ ou $[\text{PtCl}(\text{dien})] \text{Cl}$ [89, 90] (figure 1.3).
- les complexes dits « de Farrel » comportant deux platines séparés par une chaîne aliphatique. Deux types sont décrits [62, 63] :
 1. deux unités de *cis*- $\text{Pt}(\text{ammine})_2$ bifonctionnels reliées par une chaîne diammine de longueur variable (figure 1.4 a) ;
 2. complexes bis(platine) monofonctionnels cationiques (figure 1.4 b) ;

Ces deux types de complexe sont actifs *in vitro*. Les composés de type 1 sont également actifs pour les lignées cellulaires résistantes au cisplatine et *in vitro*.

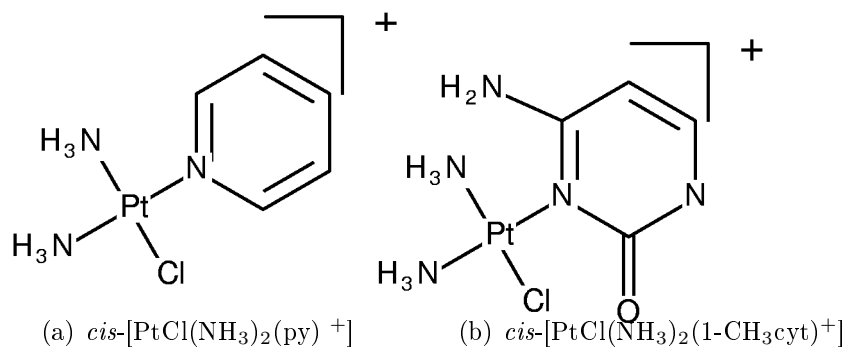


FIG. 1.3 – Les complexes de « Hollis »

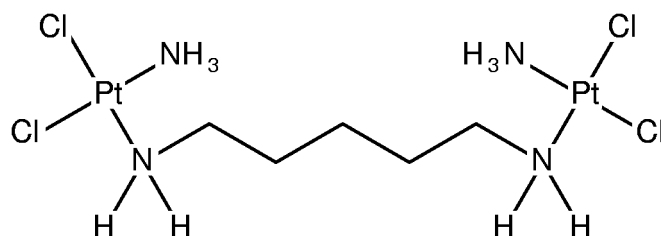
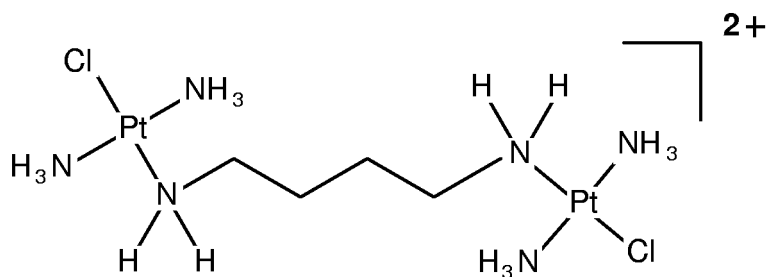
(a) $2,2/c, c[\{\text{cis-PtCl}_2(\text{NH}_3)_2\}\text{H}_2\text{N}(\text{CH}_2)_5\text{NH}_2]$ (b) $1,1/t, t[\{\text{trans-PtCl}(\text{NH}_3)_2\}_2\text{H}_2\text{N}(\text{CH}_2)_4\text{NH}_2]^{2+}$

FIG. 1.4 – Les complexes de « Farrell »

De nombreux autres types de complexes existent (notamment les complexes où les ligands partants sont en *trans*), une présentation détaillée a été réalisée par Wong et Giandomenico [207].

Type d'adduit	adduit formé (proportion)	coude (°)	déroulement (°)
intra brin (grand sillon)	GG- <i>cis</i> -Pt(NH ₃) ₂ (65%)	32-34 (EG) [15]	13 (EG) [14]
		39/55 (RX)[181] (2 conformations)	
		58 (RMN) [210]	21 (RMN) [210]
		60 (RMN) [87]	12-19 (RMN) [87]
		78 (RMN) [71]	25 (RMN) [71]
intra brin (grand sillon)	AG- <i>cis</i> -Pt(NH ₃) ₂ (25%)	32-34 (EG) [15]	13 (EG)
		55 (RMN) [65]	8-21 (RMN) [65]
intra brin (grand sillon)	GXG- <i>cis</i> -Pt(NH ₃) ₂ (6%)	32-34 (EG) [15]	23 (EG) [14]
		20 (RMN) [191]	19 (RMN) [191]
inter brins (petit sillon)	GG- <i>cis</i> -Pt(NH ₃) ₂ (4%)	45 (EG) [124]	79 (EG) [124]
		20 (RMN) [94]	87 (RMN) [94]
		40 (RMN) [140]	76 (RMN) [140]

TAB. 1.2 – Adduits majeurs provoqués par le cisplatine et caractéristiques principales (EG : électrophorèse sur gel, RMN : résonance magnétique nucléaire, RX : diffraction aux rayons X)

1.3 Action supposée du cisplatine sur les cellules cancéreuses

1.3.1 Adduits formés

La cible majeure du cisplatine est l'ADN [57], le tableau 1.2 présente les caractéristiques des adduits observés.

La valeur du coude obtenue pour les différents adduits varie significativement selon la méthode de mesure utilisée (électrophorèse sur gel, RMN, diffractions des rayons X). Le coude observé durant les simulations de dynamique moléculaire présentées dans le cadre de ce travail de thèse donnent des valeurs comprises entre celles mesurées en diffraction des rayons X et de la RMN.

1.3.2 Reconnaissance des adduits du platine

La modification conformationnelle de l'ADN par la fixation du complexe de platine induit la reconnaissance de ces lésions par des protéines [216]. En particulier les protéines possédant une séquence homologue de 80 acides aminés, appelée domaine HMG [133, 203], présentent une grande affinité pour les adduits GG-*cis*-Pt(NH₃)₂.

Les protéines à domaine HMG ont la faculté de courber l'ADN linéaire et de reconnaître des régions coudées de l'ADN. Une fois liées à l'ADN modifié par le cisplatine, ces protéines accentuent le coude dû aux adduits GG-*cis*-[Pt(NH₃)₂] et AG-*cis*-[Pt(NH₃)₂] jusqu'à une valeur de 70 à 90°[37]. Les protéines HMG se lient dans le petit sillon de l'ADN et reconnaîtraient la courbure induite par les adduits 1,2 intra-brin du cisplatine [133]. La reconnaissance de l'adduit GG-*cis*-[Pt(NH₃)₂] inter-brins du cisplatine a également été observée [101] ; l'adduit 1,3 intra-brin qui accroît la courbure de l'ADN n'est pas reconnu [148]. Pour l'adduit GTG-*cis*-[Pt(NH₃)₂], l'empilement de la thymine au dessus de la guanine 5' platinée perturbe davantage le petit sillon de l'ADN que les adduits GG-*cis*-[Pt(NH₃)₂] et AG-*cis*-[Pt(NH₃)₂][191]. Les adduits du transplatine ne sont pas reconnus [133].

Les affinités du domaine HMG isolé ou de la protéine entière sont comparables ($K_d = 10^{-6}$ à 10^{-7} M). La platination de l'ADN augmente cette affinité de 4 à 100 fois comparée à l'affinité de la protéine avec l'ADN non modifié [216]. Le domaine A de la protéine HMG1 a une affinité qui varie de plus de deux ordres de grandeur suivant les séquences étudiées, celle-ci étant maximale pour AG*G*A (l'astérisque indiquant la base platinée). Dans ce cas, l'affinité est alors augmentée d'un facteur 1000 par l'adduit GG-*cis*-[Pt(NH₃)₂] [55].

D'autres protéines qui ne possèdent pas de domaine HMG peuvent également se lier aux adduits du cisplatine [133, 143, 189]. Récemment, il a été montré que les protéines histones, protéines qui sont 10 fois plus abondantes dans le noyau que HMG1, ont une affinité 20 fois plus grande pour l'ADN platiné que HMG1 [209, 216]. Ces protéines, comme les protéines à domaine HMG, se lient à l'ADN de liaison de la chromatine et reconnaissent les coudes et le déroulement de l'ADN.

La reconnaissance des adduits du cisplatine par les protéines intervient à différents niveaux dans la modulation de l'activité antitumorale :

- les protéines peuvent servir d'éléments de reconnaissance pour la réparation, par le système d'excision-réparation du noyau (NER : *Nucleotide Excision Repair*), des nucléotides. Par exemple la protéine hRPA, chargée de la reconnaissance des lésions lors de la réparation par excision, a quatre à six fois plus d'affinité pour l'adduit GG-*cis*-[Pt(NH₃)₂] que pour l'ADN non platiné [144]. Les protéines du système de reconnaissance et de réparation des mésappariements sont déficientes dans certaines cellules résistantes au cisplatine. La reconnaissance des adduits cisplatine-ADN par les protéines de réparation semble servir de signal initiant les événements menant à l'apoptose.
- les protéines peuvent bloquer l'accès des adduits aux protéines de réparation [144]. L'inhibition de la réparation des adduits GG-*cis*-[Pt(NH₃)₂] et AG-*cis*-[Pt(NH₃)₂] est provoquée par la fixation des protéines HMG1 et tsHMG empêchant la pro-

téine hRPA de s'y fixer.

- les adduits du cisplatine entreraient en compétition avec le site naturel de liaison et bloqueraient ainsi sa fonction normale. Par exemple le facteur de transcription ribosomal hUBF à domaine HMG A, *in vitro*, une affinité du même ordre de grandeur pour son site naturel de liaison que pour l'adduit GG-*cis*-[Pt(NH₃)₂]. A des concentrations en cisplatine plus faibles que dans les cellules traitées, les adduits du cisplatine peuvent détourner hUBF de sa cible naturelle [188].

1.3.3 Inhibition de la réplication

In vivo et *in vitro*, le cisplatine inhibe la répplication de l'ADN [22, 149]. Les diadduits du cisplatine bloquent la synthèse de l'ADN avec des efficacités différentes. La lésion majoritaire du cisplatine GG-*cis*-[Pt(NH₃)₂] est la plus efficace [43]. Les monoadduits formés par [Pt(dien)Cl]⁺ ne bloquent pas la synthèse de l'ADN [85, 149]. Toutefois, l'activité des monoadduits semble dépendre de la polymérase étudiée [88].

L'inhibition de la réplication ne semble pas la cause principale de la mort cellulaire. En effet, le transplatine, isomère cliniquement inactif du cisplatine, inhibe la réplication de l'ADN avec le même taux de lésions sur l'ADN que le cisplatine [39]. De plus, la dose nécessaire pour bloquer la réplication est supérieure à la dose qui induit la mort cellulaire dans les cellules leucémiques L210 [167].

Les cellules traitées par le cisplatine arrêtent leur cycle cellulaire en phase G2 donc après la réplication de l'ADN et avant la mitose [177, 178].

1.3.4 Inibition de la transcription

Les diadduits du cisplatine sur l'ADN inhibent la transcription [44, 45, 46], contrairement aux monoadduits du [Pt(dien)Cl]⁺ qui n'induisent qu'une inhibition partielle de la transcription. De même que les adduits du cisplatine, ceux du transplatine inhibent la transcription mais de façon moins efficace. Le taux de transcription translésionnelle (au-delà de l'adduit) de l'ARN polymérase est beaucoup plus élevé après traitement par le transplatine qu'après traitement par le cisplatine (60-76 % et 0-16 % respectivement). L'inhibition de la transcription pourrait donc contribuer à la différence d'activité entre ces deux isomères.

Il est cependant à noter que les gènes dont le taux de transcription est élevé sont les plus rapidement réparés [79], de sorte qu'il faut une quantité de cisplatine plus importante pour inhiber la transcription que pour inhiber la réplication.

1.3.5 Réparation des adduits du cisplatine

Sur les leucocytes de patients traités par le cisplatine, plus les adduits cisplatine-ADN sont réparés, plus la réponse clinique est faible [213]. De même, les cellules déficientes en système de réparation sont plus sensibles au traitement par le cisplatine que les cellules sauvages [52] ; et les cellules résistantes à la drogue ont une capacité de réparation supérieure à celle des cellules sauvages.

Tous les adduits du cisplatine ne sont pas réparés avec la même efficacité. Les adduits GXG-*cis*-[Pt(NH₃)₂] sont réparés 15 à 20 fois plus efficacement que les adduits GG-*cis*-[Pt(NH₃)₂] et AG-*cis*-[Pt(NH₃)₂]. A partir d'extraits de cellules humaines, l'adduit GG-*cis*-[Pt(NH₃)₂] n'est pas réparé par le système de réparation par excision des nucléotides. Une fois l'adduit excisé par le système UvrABC de *E. coli*, et après une étape de ligation, la réparation est complète. Les différentes études ont montré que l'ordre des vitesses de réparation des adduits du cisplatine est le suivant : GXG-*cis*-[Pt(NH₃)₂] > AG-*cis*-[Pt(NH₃)₂] > GG-*cis*-[Pt(NH₃)₂]. Ces résultats impliquent que les enzymes humaines du système de réparation excisent difficilement l'adduit GG-*cis*-[Pt(NH₃)₂]. Il est à noter que les études *in vitro* ont montré que le système de réparation par excision des nucléotides est le système impliqué dans la réparation des adduits cisplatine-ADN.

Si les guanines de l'adduit GG-*cis*-[Pt(NH₃)₂] sont mésappariées avec une ou deux thymines, la réparation est augmentée de façon importante. La différence structurale entre les adduits 1,2 et 1,3-cisplatine semble donc responsable de la différence de réparation entre les adduits.

La capacité de réparation des lésions du cisplatine sur l'ADN varie toutefois en fonction du tissu dont sont extraites les protéines, mais n'est pas liée au profil de toxicité du cisplatine.

1.3.6 Propriétés mutagènes des adduits du cisplatine

A long terme, le cisplatine provoque des mutations impliquées dans le développement de tumeurs secondaires. La réplication translésionnelle des adduits cisplatine-ADN est une des source des mutations.

Les adduits GG-*cis*-[Pt(NH₃)₂] et AG-*cis*-[Pt(NH₃)₂] produisent respectivement majoritairement des transversions de G_{5'} en T et de A_{5'} en T [21, 24, 212]. L'adduit AG-*cis*-[Pt(NH₃)₂] est environ 5 fois plus mutagène que l'adduit GG-*cis*-[Pt(NH₃)₂]. Aucune mutation spécifique n'a été observée au niveau de l'adduit GXG-*cis*-[Pt(NH₃)₂] alors qu'il bloque moins efficacement les ADN polymérases que les adduits 1,2-*cis*-Pt. La transversion spécifique de la base 5' des adduits 1,2-*cis*-Pt peut être corrélée à la plus grande déformation de la base du côté 5'. Cette observation doit être pondérée

par la faible mutagénicité des adduits 1,3-*cis*-Pt alors que ceux-ci déforment davantage la double hélice.

Une proportion importante des mutations induites par le cisplatine se traduit aussi par la délétion de fragments d'ADN [23, 40].

1.3.7 Résistance au cisplatine

La résistance cellulaire acquise au cisplatine qui apparaît après une réponse initiale de la tumeur à la chimiothérapie est la cause majeure des échecs du traitement des cancers par le cisplatine.

Plusieurs mécanismes sont impliqués dans ce phénomène :

- l'accumulation du cisplatine dans la cellule peut diminuer de 50%. Cependant la diminution de la quantité de cisplatine à l'intérieur de la cellule peut être faible, même pour des niveaux de résistance élevés [103, 146].
- cette décroissance s'accompagne d'une augmentation du taux intracellulaire de composés soufrés tels que le glutathion et les métallothionéines, ainsi que de l'activité de la GSH transférase. La pompe à GS-X est suractivée ce qui entraîne le transport quatre fois plus important des complexes glutathion-cisplatine à travers la membrane chez les cellules résistantes au cisplatine par rapport aux cellules sensibles. Après diminution du glutathion intracellulaire, la sensibilité de la cellule au cisplatine n'est pas recouverte [1, 13, 74, 136].
- la résistance au cisplatine est liée à l'augmentation des facteurs de reconnaissance des lésions de l'ADN et du taux de réparation des adduits; alors que la faible résistance au cisplatine s'accompagne de la perte de reconnaissance des adduits du cisplatine sur l'ADN par le système de réparation de mésappariements de l'ADN [20, 97, 98, 110, 130, 131, 141, 214].

Chapitre 2

L'ADN

Depuis la publication de son organisation tridimensionnelle [199], la description des différents assemblages de la molécule a été normalisée par l'IUPAC/IUB. Voici donc les conventions et la nomenclature qui ont été adoptées pour les acides nucléiques puis brièvement les structures classiques de l'ADN.

Les acides nucléiques (ADN et ARN¹) sont des structures polynucléotidiques. L'unité nucléotidique est composée d'un sucre furanose (le β -D-2'-désoxyribose pour l'ADN et le β -D-ribose² pour l'ARN) phosphorylé en position 5' ou 3' (nucléotide du type 5'-mono-phosphate ou 3'-mono-phosphate) et substitué en position 1' par une base azotée. Les nucléotides sont reliés par des liaisons covalentes entre les phosphates et les sucres successifs. L'enchaînement sucre-phosphate constitue le squelette phosphodiester chargé négativement. Par convention, les polynucléotides sont tracés et lus dans le sens 5' \longrightarrow 3'. L'enroulement autour d'un même axe de deux chaînes polynucléotidiques antiparallèles et complémentaires au niveau des bases forme la double hélice de l'ADN [199]. Les deux brins antiparallèles du squelette phosphodiester permettent de définir deux sillons dans la double hélice (petit et grand) qui ont des caractéristiques différentes dans les diverses formes de l'ADN [figure 2.2].

2.1 Les bases

Les cinq bases des acides nucléiques sont regroupées en fonction de leur nature chimique : la guanine (G) et l'adénine (A) sont des purines (R), la cytosine (C), la thymine

¹L'ARN se distingue de l'ADN par sa composition en bases et par son groupement C2'-hydroxy. Sa structure secondaire se compose de « tiges » ou zones appariées formant une structure semblable à celle de l'ADN, et de « boucles » ou zones non appariées entre deux « tiges ». Dans la structure tertiaire, on observe des interactions qui impliquent des nucléotides distants dans la séquence et qui résultent en des conformations très variées et souvent très différentes de celles de l'ADN. Les ARN peuvent avoir une activité enzymatique (ribozymes). Les ribozymes sont impliqués dans le métabolisme, notamment au niveau de la transcription [8, 27, 153].

²Le terme ribose vient de *Rockefeller Institut for Biochemistry* (rib) et de ose (sucre)

(T) et l'uracile (U) sont des pyrimidines (Y). Dans l'ARN, l'uracile, qui remplace la thymine, diffère de celle-ci par l'absence de groupement méthyle en C5 (figure 2.1).

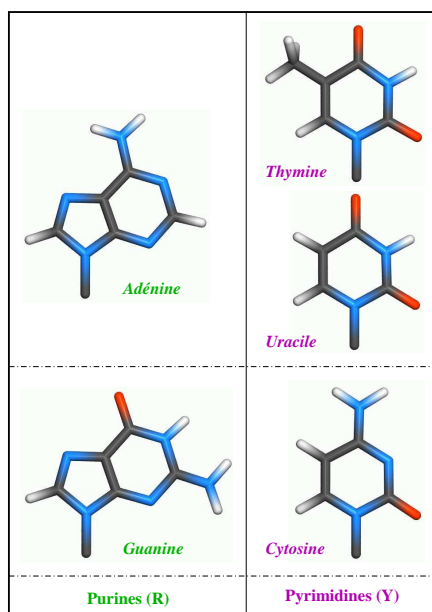


FIG. 2.1 – Les bases nucléotidiques.

bleu : azote, gris : carbone, blanc : hydrogène, rouge : oxygène.

La forme canonique de l'ADN est générée par l'appariement spécifique entre les bases de chaque brin formant la double hélice. Les bases sont planes de sorte qu'elles sont empilées les unes sur les autres le long du brin (*stacking*). Les interactions entre les systèmes conjugués des cycles aromatiques des bases d'un même brin et des brins complémentaires contribuent fortement à stabiliser la structure hélicoïdale. L'énergie d'empilement provient essentiellement des interactions de type Van Der Waals. Dans la double hélice, l'enthalpie de *stacking* entre deux paires adjacentes est estimée entre -5 et -12 Kcal/mol [126].

L'appariement le plus fréquemment rencontré est celui du type Watson-Crick [199] où une purine est associée à une pyrimidine (C·G, A·U, ou A·T) (figure 2.2). Les nombreux groupes donneurs et accepteurs de liaisons hydrogène des quatre bases se prêtent à d'autres assemblages qui, comme les paires Watson-Crick, sont approximativement coplanaires (figure 2.3). Certains associent A avec T, ou G avec C, comme les paires

Hoogsteen [91], mais on trouve aussi des paires G·T (*wobble*), G·A, G·G, C⁺·C³, T·T, A·A, ainsi que des triplets et des quadruplets de bases.

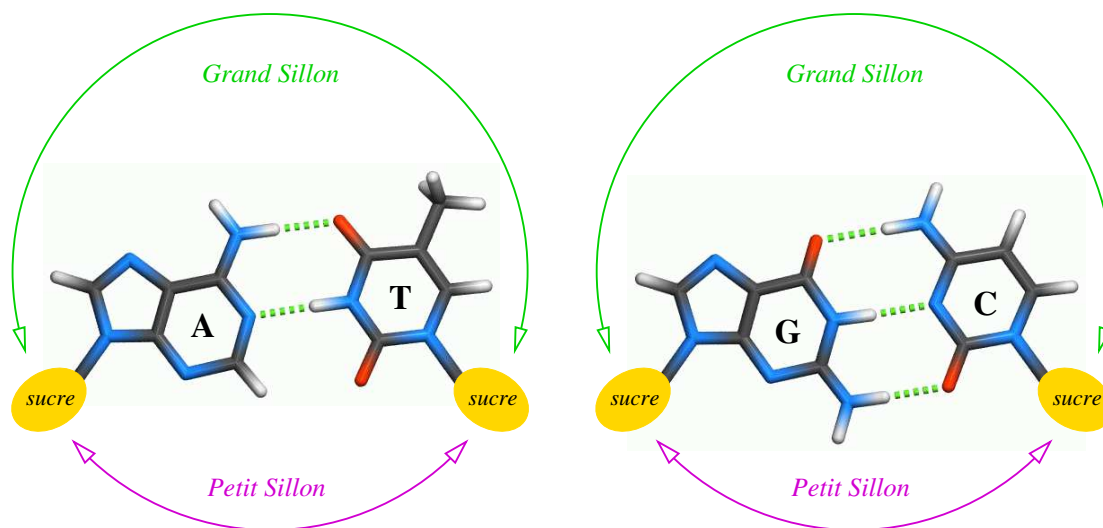


FIG. 2.2 – Appariements Watson-Crick et définition des sillons

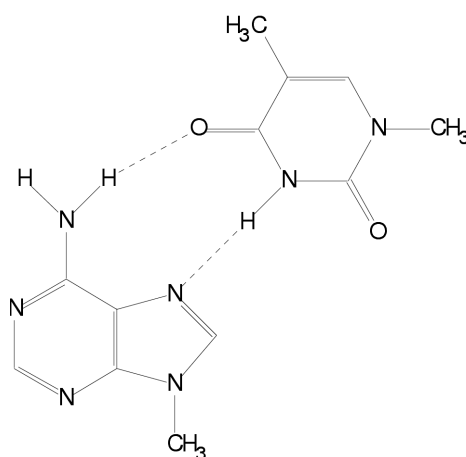


FIG. 2.3 – Appariement selon Hoogsteen entre l'adénine et la thymine, dans la structure cristalline formée par la paire 9-méthyladénine/1-méthylthymine [194].

L'interaction entre les bases d'une paire est assurée par des liaisons hydrogène constituées majoritairement d'interactions électrostatiques. Il y a deux liaisons pour les paires Watson-Crick A·U et A·T et trois pour la paire G·C. Ainsi l'appariement entre une guanine et une cytosine est plus favorable que celui entre une thymine et une adénine. Il faut également noter que dans les solvants polaires, les molécules d'eau rentrent en compétition avec les bases pour la formation des liaisons hydrogène avec les bases complémentaires, réduisant ainsi l'énergie libre d'appariement entre les paires

³ C⁺ : cytosine protonnée sur la position imino (N3).

de bases. Les expériences thermodynamiques effectuées dans le chloroforme et dans le diméthyl-sulfoxyde estiment aux alentours de -2 à -5 Kcal/mol l'enthalpie d'appariement entre deux bases [109, 138] et de récents calcul d'énergie libre estiment l'énergie libre d'appariement des paires A·T vers -4.3 Kcal/mol et vers -5.8 Kcal/mol pour les paires G·C [6, 179]. Ces calculs montrent que les interactions sont certainement plus faibles que celles du *stacking*. Ainsi la spécificité de l'appariement est assurée par les liaisons entre les bases formant une paire et l'empilement de ces paires est l'origine principale de la stabilité de l'hélice. La répulsion entre les phosphates anioniques tend à déstabiliser la structure et la présence des contre-ions est nécessaire pour éviter la dénaturation.

2.2 Le squelette phosphodiester

La conformation du squelette phosphodiester nécessite la définition de six angles de torsion pour chaque nucléotide. Ces torsions correspondent aux liaisons simples qui se succèdent le long du squelette dans le sens 5'→3'. Elles sont notées par les lettres α , β , γ , δ , ε , ζ avec $\alpha(\text{P-O5'})$ et $\zeta(\text{O3'-P})$ (figure 2.4).

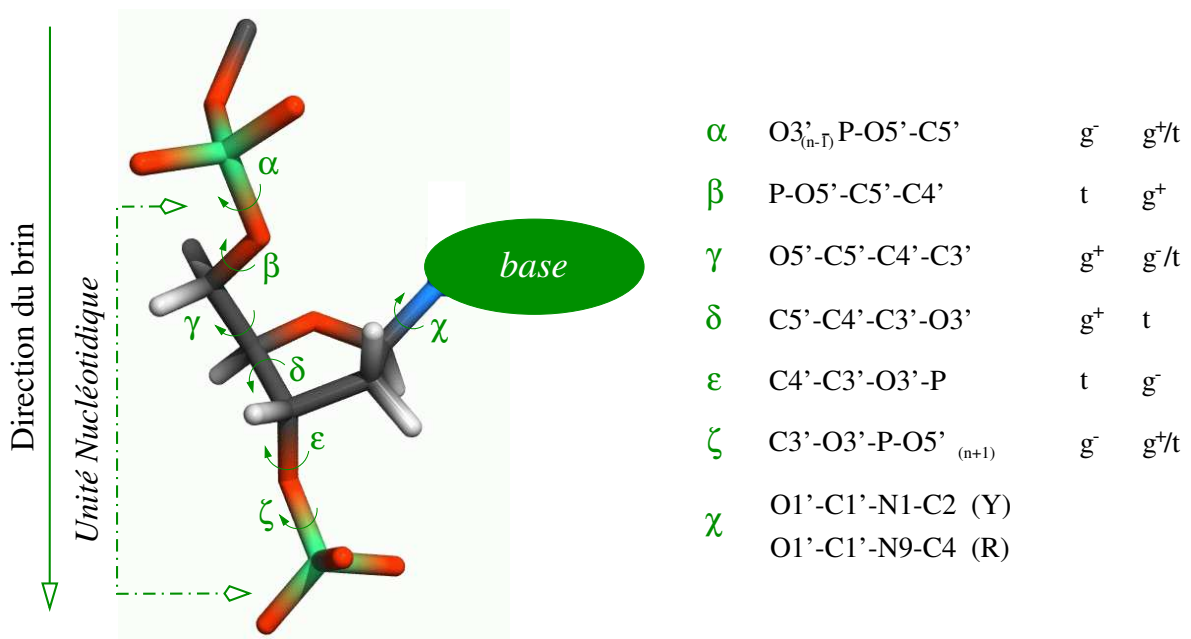


FIG. 2.4 – Le squelette phosphodiester et les angles le définissant.

Grâce aux nombreux degrés de liberté autour des liaisons simples, le squelette phosphodiester est assez flexible. Cependant, l'environnement stérique définit autour de chaque liaison des rotations autorisées et entraîne des conformations privilégiées pour chacun des angles de torsion. Ces conformations sont appelées « gauche moins » (g^-), « trans » (t), et « gauche plus » (g^+) et leurs valeurs sont proches de -60° , 180°

et $+60^\circ$ respectivement. Les variations des angles de torsion s'effectuent souvent de manière corrélée. Ainsi dans l'ADN B, la position du groupement phosphate peut être amenée plus près du petit sillon de la double hélice en changeant la conformation des angles ε et ζ de tg^- à g^-t (transition BI/BII) [76, 154]. De même, dans l'ADN A, α et γ varient souvent ensemble et passent de la conformation g^-g^- à la conformation tt [80, 154]. Cette transition couplée s'accompagne généralement d'une modification de l'angle β , qui passe de t à g^+ .

L'angle δ appartient au cycle du sucre et prend de ce fait des valeurs variant dans une échelle plus restreinte que les autres angles de torsions, soit entre 70° et 170° . Notons pour finir que les barrières de potentiel de tous ces angles de torsion peuvent varier de 0,5 à 5 Kcal/mol environ selon leurs substituants [115].

2.3 Le sucre furanose

Les sucres furanoses des acides nucléiques sont des pentoses dont le cycle n'est pas plan mais plissé. Le plissement est dit « enveloppe » (E) lorsqu'un seul des atomes du cycle sort du plan formé par les quatre autres. Si deux atomes sortent du plan formé par les trois autres, le plissement est « *twist* » (T) (figure 2.5). Il y a donc vingt conformations dont le plissement est maximal : cinq conformations E « vers le bas » notées ${}_iE$ (où i désigne l'atome déplacé du côté opposé à l'atome C5' par rapport au plan cyclique), cinq conformations E « vers le haut » (où i désigne l'atome déplacé du côté de l'atome C5' par rapport au plan cyclique) notées iE , les cinq ${}_iT$ et les cinq iT [202].

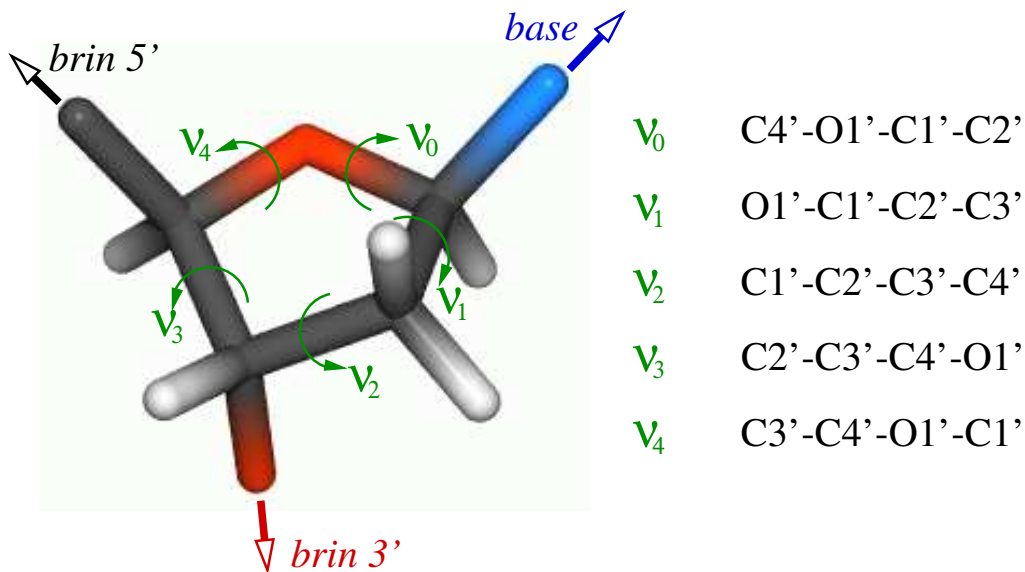


FIG. 2.5 – Le sucre furanose et les angles qui le décrivent.

Dans un cycle à cinq atomes dont les longueurs de liaison sont fixes, cinq angles de torsions décrivent la conformation (figure 2.5). Il y a quatre degrés de liberté conformationnelle : cinq atomes soit quinze coordonnées cartésiennes, auxquels il faut retrancher six degrés de liberté pour les translations et les rotations correspondant au positionnement dans l'espace, et cinq degrés de liberté correspondant aux longueurs de liaisons fixes.

Le concept conventionnel de pseudo-rotation [3] décrit le sucre comme une vague stationnaire à l'aide de deux de ces variables seulement : l'amplitude qui indique le degré de non-planarité et la phase qui indique la position et le signe du plissement. Les deux autres variables qui doivent compléter la description du cycle correspondent à des réajustements mineurs des angles de valence endo-cycliques [128, 127]. La phase P et l'amplitude m d'un sucre sont définis en fonction des angles de torsion ν par les relations suivantes :

$$P = \tan^{-1}(B/A) \text{ et } m = (A^2 + B^2)$$

avec

$$A = \frac{2}{5} \sum_{j=0,4} \nu_j \cos\left(\frac{4\pi j}{5}\right)$$

$$B = -\frac{2}{5} \sum_{j=0,4} \nu_j \sin\left(\frac{4\pi j}{5}\right)$$

Le cercle de pseudo-rotation (figure 2.6) représente les vingt conformations possibles que peut adopter un pentose. Toutes les conformations ne sont pas autorisées dans le nucléotide car le sucre est asymétriquement substitué. Les conformations les plus stables sont les formes C3'-endo (sucre « nord », N) et C2'-endo (sucre « sud », S) (figure 2.7). La RMN [28] et la cristallographie [215] indiquent également que la région centrée autour de O1'-endo constitue un troisième domaine stable (sucre « est », E). Le profil énergétique d'un nucléotide en fonction de la phase de son sucre montre les régions stables C2'-endo et C3'-endo et indique également que la transition de l'une vers l'autre se fera par la conformation O1'-endo et non O1'-exo, du fait de l'encombrement stérique généré par la base en C1'.

Un dernier angle de torsion complète la description du squelette phosphodiester : l'angle glycosidique χ (O1'-C1'-N1-C2 pour les pyrimidines et O1'-C1'-N9-C4 pour les purines) qui relie la base azotée au sucre. Il peut adopter deux conformations notées *syn* et *anti*. Dans les acides nucléiques hélicoïdaux, la conformation *anti* est dominante car elle minimise l'encombrement stérique entre les atomes de la base et ceux du sucre. La conformation *syn* est observée dans des structures particulières comme l'ADN Z, généralement pour des purines où les gênes stériques sont moins importantes (figure 2.8).

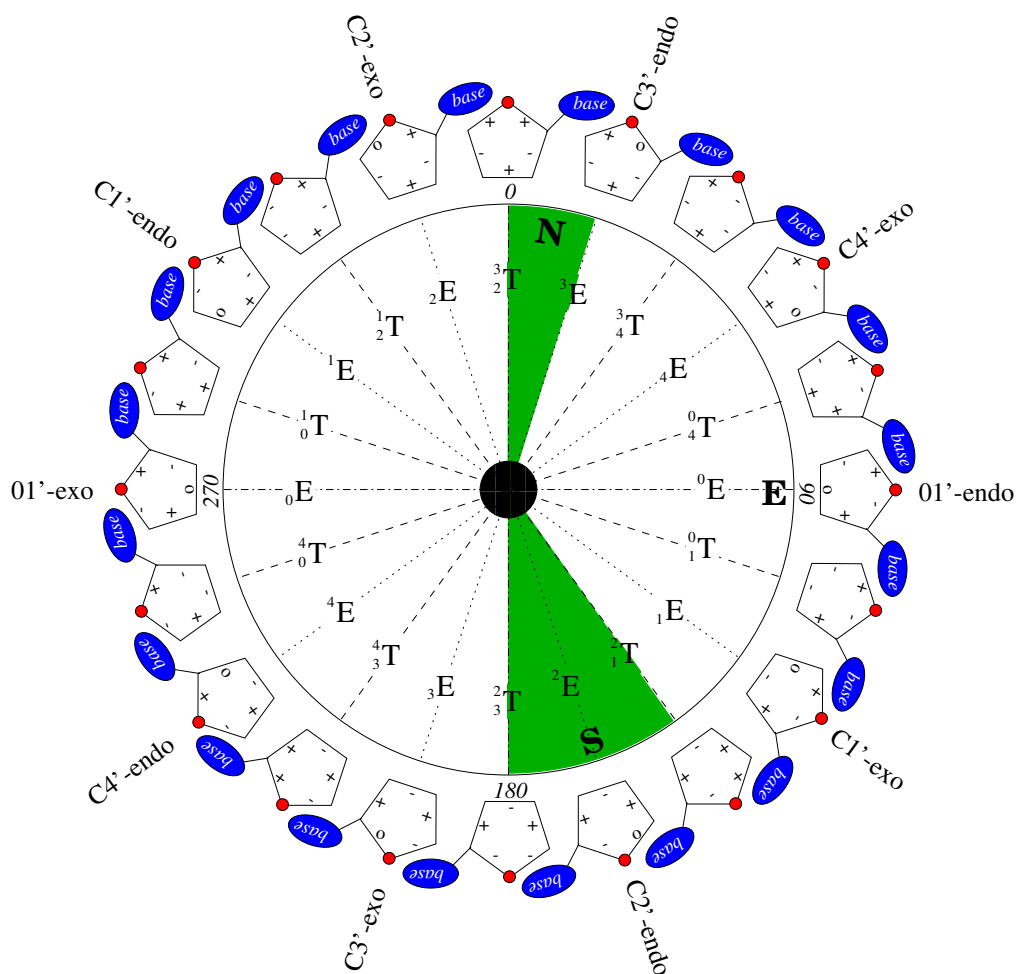


FIG. 2.6 – Cercle de pseudo-rotation. Les signes des cinq angles dièdres endo-cycliques sont indiqués à l'intérieur du cycle furanose [3]

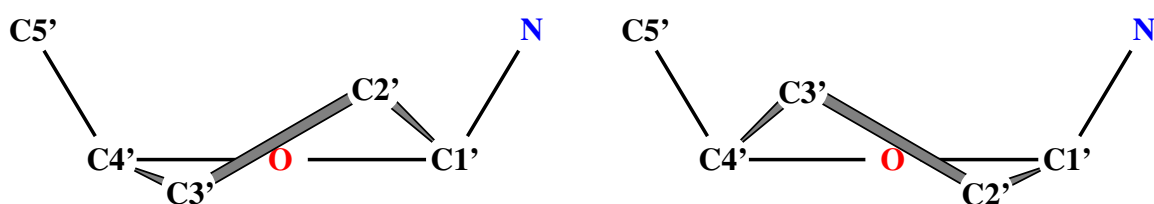


FIG. 2.7 – Représentation schématique des plissements C2'-endo (gauche) et C3'-endo (droite)

2.4 Les paramètres hélicoïdaux

Le plissement des sucres et les angles de torsion du brin phosphodiester permettent de caractériser complètement la structure interne de chaque brin d'un acide nucléique. Toutefois, l'utilisation de paramètres qui décrivent leur conformation hélicoïdale donne une meilleure image de leur structure globale. De tels paramètres ont été utilisés dès

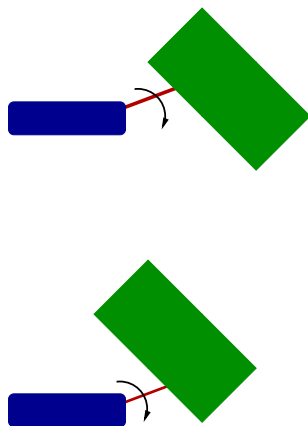


FIG. 2.8 – Configuration syn et anti de l'angle glycosidique χ . Le sucre et la base sont schématisés en bleu et en vert, respectivement.

que les premières données sur la conformation de l'ADN à partir des expériences sur les fibres d'ADN furent accessibles. Le nombre de paramètres augmenta avec la découverte de la multitude d'irrégularités structurales de la double hélice, que l'on a pu observer dès l'obtention des mono-cristaux d'ADN. Lors de la convention de Cambridge en 1988 [Dickerson *et al.*, 1989], un ensemble de paramètres hélicoïdaux standards permettant de décrire complètement la structure des acides nucléiques a été défini.

Ces paramètres sont répartis en trois groupes :

1. Les quatre paramètres de positionnement d'une base par rapport à l'axe hélicoïdal (deux rotations et deux translations) (figure 2.9).
2. Six paramètres intra-paire de bases pour positionner les deux bases de la paire l'une par rapport à l'autre (trois translations et trois rotations) (figure 2.10).
3. Six paramètres inter-paires de bases pour orienter et positionner deux paires de bases successives (trois translations et trois rotations) (figure 2.11).

Le sens positif des angles de rotation est donné par la « règle de la main droite » autour des axes associés. Ces paramètres ne présentent pas tous le même intérêt : certains en effet ont une variabilité réduite dans les conformations standards (*shear*, *stretch*, *stagger* et *opening*).

La manière dont doivent être calculés les paramètres n'a pas été définie à Cambridge et aujourd'hui il existe deux approches différentes : l'approche dite « globale » découle des travaux originaux sur les hélices régulières, elle définit un axe par rapport auquel sont calculés les paramètres hélicoïdaux. Dans les structures déformées, un axe global droit n'est pas approprié mais la définition d'un axe global courbe permet de donner une image des grandes déformations possibles de la double hélice [112]. Au contraire,

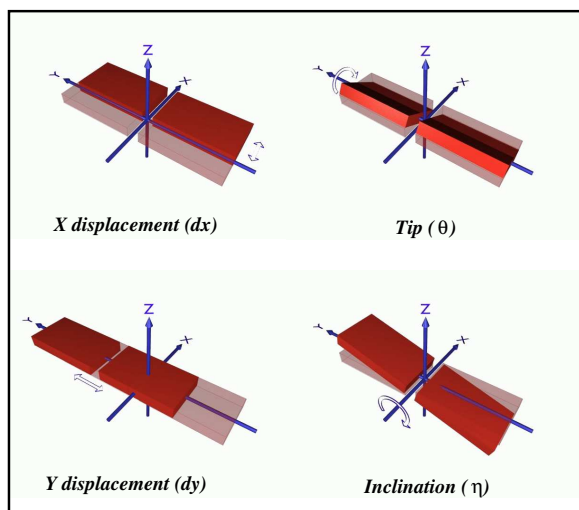


FIG. 2.9 – Paramètres hélicoïdaux

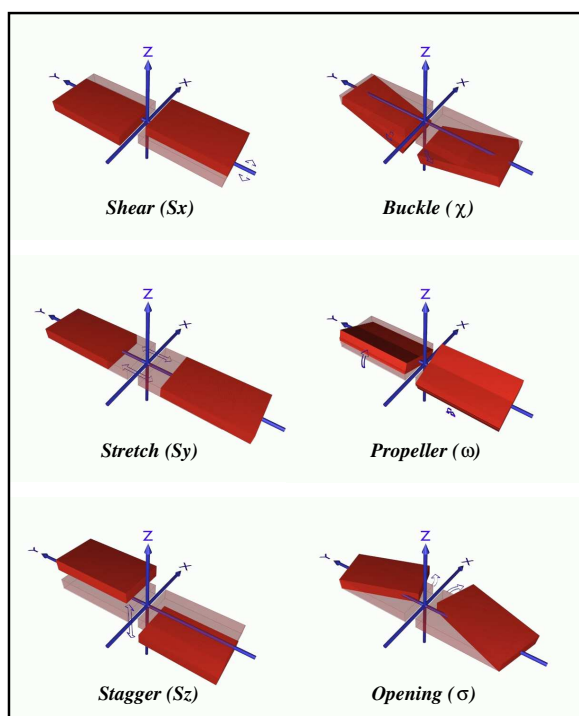


FIG. 2.10 – Paramètres intra-paires de bases

l'approche « locale » s'attache à décrire la jonction entre paires de bases adjacentes et ne se préoccupe pas de définir un axe global [114].

Finalement, il est souvent intéressant de mesurer la géométrie des sillons formés entre les brins phosphodiester. Le petit sillon correspond par convention au côté des liaisons glycosidiques, le grand sillon correspond à l'autre côté (figure 2.2). Ainsi, suivant la conformation de l'ADN, les dimensions des sillons peuvent varier beaucoup. La

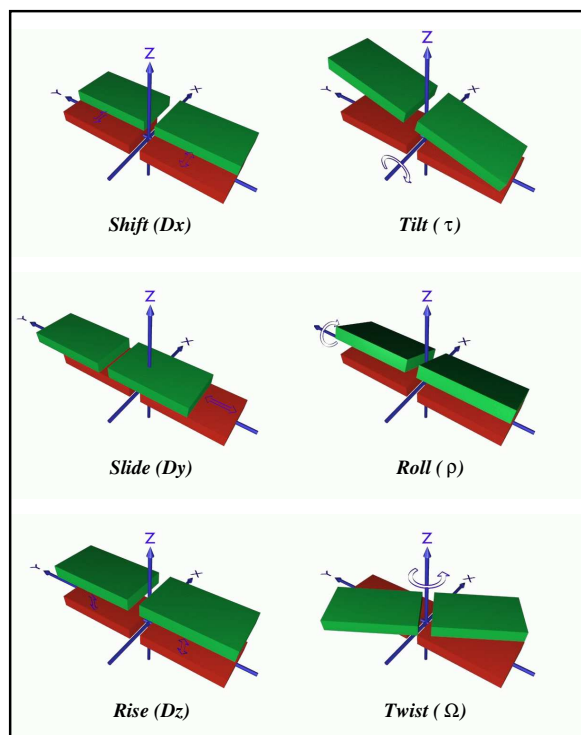


FIG. 2.11 – Paramètres inter-paires de bases

largeur des sillons peut être calculée de deux façons différentes : comme la distance la plus courte entre deux atomes de phosphore à travers le sillon, ou de façon continue le long du sillon [168].

2.5 Polymorphisme de l'ADN

La cristallographie reste la méthode la plus apte à fournir des informations au niveau atomique sur les structures des acides nucléiques [50, 86, 171, 186]. Elles sont rassemblées dans les bases de données standards (la *Protein Data Bank* « PDB » [17, 19] ou la *Nucleic Acids Data Bank* « NDB » [18]). Elles contiennent plus de cent structures d'ADN résolues à moins de 2 Å, une résolution d'au moins 1.4 Å étant nécessaire pour déterminer avec précision les paramètres hélicoïdaux [155]. On peut émettre des réserves quant aux informations apportées par la cristallographie puisque l'environnement cristallin peut induire des déformations importantes de la double hélice qu'elle ne subit probablement pas dans les conditions physiologiques. Cependant, il faut noter que l'ADN dans la cellule se trouve en contact étroit avec d'autres molécules (protéine ou ADN), donc dans un environnement de faible constante diélectrique, de haute force ionique, dans une situation de déshydratation partielle, et sa structure n'est peut être pas si éloignée de celle des ADN cristallisés.

La RMN permet de déterminer des structures moyennées dans le temps d'oligomères en solution [96]. On peut ainsi distinguer les différents allomorphes de l'ADN et les perturbations structurales majeures induites par des paires de bases non complémentaires comme les mésappariements (*mismatch*) ou par des zones de non appariement (*bulge*). Elle permet, comme la cristallographie, d'observer des variations structurales de l'ADN qui dépendent de sa séquence [35, 132]. Toutefois, les résultats obtenus par la RMN sont encore dépendants des techniques de modélisation utilisées pour affiner les modèles de structure.

La structure globale de la double hélice d'ADN est dépendante de son environnement : force ionique, degré d'hydratation, température, nature du solvant [171] et aussi de sa composition en bases. Les nombreuses études de diffraction des rayons X sur des fibres et des cristaux d'oligonucléotides d'une part, et les expériences de RMN d'autre part classent les structures de l'ADN en trois grandes familles allomorphes : A, B et Z. Ces trois structures sont visuellement reconnaissables : à nombre égal de paires de bases, la double hélice est longue et élancée dans la conformation B (figure 2.12), courte et large dans la conformation A (figure 2.13) et plus longue et plus fine dans la conformation Z (figure 2.14). Les valeurs caractéristiques moyennes des paramètres structuraux de ces différentes formes sont regroupées dans le tableau 2.1. Ces conformations diffèrent par l'empilement des paires de bases et par leur position par rapport à l'axe hélicoïdal.

2.5.1 Forme B

La forme B de l'ADN est dominante lorsque l'hygrométrie est supérieure à 90% [66, 206]. L'enroulement de l'hélice autour de son axe étant « main droite », on parle d'une hélice droite. Les paires de bases se situent au centre de la double hélice, et sont globalement perpendiculaires à l'axe hélicoïdal. Il y a environ dix paires de bases par tour, espacées de 3.4 Å le long de l'axe (figure 2.12). Les sucres sont « sud », les torsions glycosidiques sont *anti* et les groupements phosphate sont en forme BI ($\varepsilon - \zeta = -90^\circ$ contrastant avec une valeur d'environ $+90^\circ$ pour la forme BII). Les groupements phosphate sont diamétralement opposés et leurs oxygènes anioniques pointent vers l'extérieur de l'hélice. Ces derniers sont stabilisés par des interactions avec les molécules d'eau et les contre-ions positifs, réduisant ainsi la répulsion électrostatique entre les deux brins. Du fait que les deux liaisons glycosidiques d'une paire de bases sont du même côté, les deux sillons n'ont pas la même dimension. Ainsi le petit sillon est étroit et le grand sillon est large, tandis que leurs profondeurs sont comparables. Cette différence est très importante pour les interactions avec les protéines ou d'autres molécules : chaque sillon forme un environnement local différent du point de vue électrostatique et stérique.

Paramètre	ADN A	ADN B	ADN Z
paramètres hélicoïdaux globaux			
<i>Xdisp</i>	-5,3	0,0	1,2
<i>Ydisp</i>	0,0	0,0	2,0/2,0
<i>Inc</i>	20,7	1,5	14,5
<i>Tip</i>	0,0	0,0	178,5/-178,5
<i>Buckle</i>	0,0	0,0	4,9/-4,9
<i>Propeller Twist</i>	-7,5	-13,3	1,1
<i>Slide</i>	0,0	0,0	3,9/-3,9
<i>Rise</i>	2,6	3,4	4,7/2,7
<i>Tilt</i>	0,0	0,0	0,0
<i>Roll</i>	0,0	0,0	-2,9/2,9
<i>Twist</i>	32,7	36,0	-8,9/-49,2
angles de torsion du squelette			
α	-74,8	-40,7	71,7/-137,4
β	-179,1	135,6	-176,0/-169,1
γ	58,9	37,4	175,5/60,0
ε	-155,0	-133,2	-92,2/-101,8
ζ	-67,1	-156,9	-53,3/76,6
χ	-158,9	-101,9	63,3/-147,4
description du sucre			
Phase	18,3	154,8	13,2/163,4
Amplitude	41,6	39,7	38,9/17,1

TAB. 2.1 – Paramètres structuraux des trois familles de l'ADN. Les translations sont en Å et les rotations en degrés. Pour l'ADN Z, les valeurs pour les paires YpR puis pour les paires RpY sont données. Elles reflètent la symétrie dinucléotidique de cette forme [81].

Selon certains auteurs, au niveau du petit sillon des paires de bases A·T, deux couches successives de molécules d'eau formeraient une « épine dorsale » d'hydratation : celles de la première couche reliraient deux accepteurs de liaisons hydrogène, tout en étant elles même reliées entre elles par les molécules de la deuxième couche [54]. L'étude récente de la structure du dodécamère de Dickerson résolue à 1.4 Å suggère également la présence d'ions monovalents, notamment au niveau des séquences du type A_nT_n [174, 175] dans la première couche de cette « épine dorsale » d'hydratation.

2.5.2 Forme A

La forme A est induite lorsque l'hygrométrie est inférieure à 75%, soit dans l'environnement cristallin : soit dans les solutions mixtes alcool/eau [104]. Elle est également favorisée par des séquences riches en paires de bases GC, mais reste généralement moins stable que l'ADN B en solution aqueuse [135, 198]. Ainsi, bien qu'il soit facile d'obtenir des cristaux d'ADN en forme A, ils contiennent parfois un mélange des formes A et B

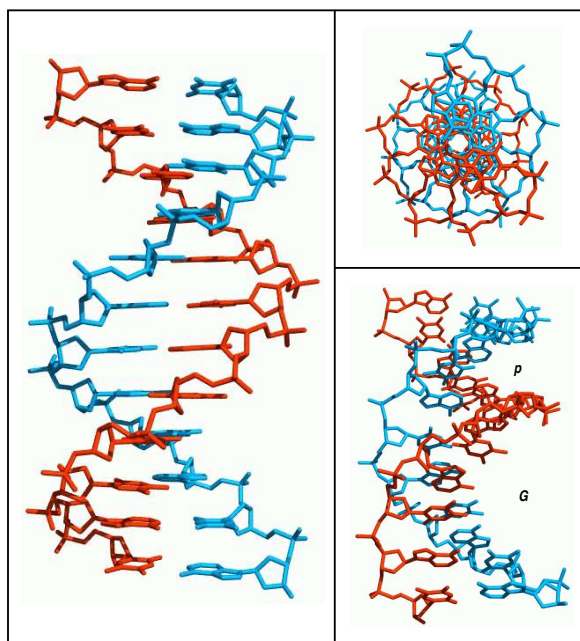


FIG. 2.12 – ADN B. P : petit sillon, G : grand sillon.

[51, 119].

L'ADN dans sa forme A (figure 2.13) est une hélice droite, son diamètre est plus grand que celui de l'ADN B. Le sucre passe de la forme C2'-endo à la forme C3'-endo. L'angle glycosidique, le *Twist* et le *Rise* sont réduits. Le pas de l'hélice est de 28 Å et de onze paires de bases par tour. Les paires de bases sont déplacées vers le petit sillon et ne sont plus perpendiculaires à l'axe de l'hélice mais inclinées (le *Tilt* est de 20°). Le déplacement des paires de bases a une conséquence sur l'aspect des sillons : par rapport à la forme B, le grand sillon est étroit et très profond, le petit sillon est très élargi et peu profond. Ces changements structuraux induisent le rapprochement des groupements phosphate vers le grand sillon : le potentiel électro-négatif aux abords de ce dernier est ainsi augmenté, ce qui peut influencer la fixation d'un éventuel ligand. Les oxygènes anioniques se retrouvent face à face et sont fortement hydratés. Ils peuvent ainsi former entre eux, du fait de leur proximité, des liaisons inter et intra-brin par l'intermédiaire des molécules d'eau [165].

On observe une moins grande variation de la structure en fonction de la séquence dans la forme A que dans la forme B [180]. Par exemple, les phases des sucres C3'-endo sont beaucoup plus confinées autour de leur valeur canonique (18°) que ne le sont les sucres C2'-endo (162°) dans la forme B. L'ADN A posséderait également une « épine dorsale » d'hydratation dans le grand sillon [64].

Les transitions B→A n'induisent que des réarrangements limités du squelette et de la géométrie des bases (tableau 2.1). Elles peuvent intervenir de façon concertée et sans

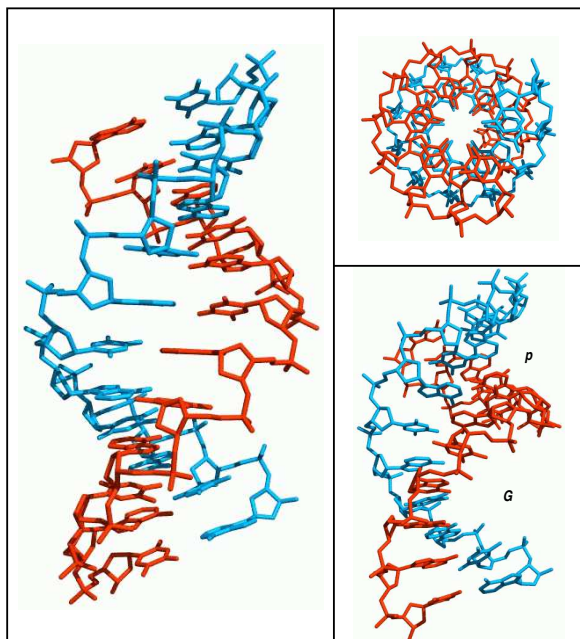


FIG. 2.13 – ADN A

à-coup dans toute la molécule d'ADN. Elles passent par des intermédiaires hybrides dont les deux brins peuvent contenir un mélange de sucres en conformation C3'-endo et C2'-endo [5, 157]. Les jonctions B/A se forment avec un faible coût énergétique puisque les bases y conservent un bon empilement et que le squelette phosphodiester y est peu étiré. En outre, elles entraînent une petite courbure de l'axe hélicoïdal (d'environ 15°) préférentielle à l'inclinaison des bases [111]. En cristallographie, des formes intermédiaires entre A et B sont observées localement entre deux fragments d'ADN-B. Les complexes ADN-Zif268⁴ [145] et ADN-TBP⁵ [77] illustrent ce phénomène. On peut ainsi penser que cette forme intermédiaire intervient dans le fonctionnement de l'ADN dans la cellule.

2.5.3 Forme Z

La forme Z a été détectée initialement grâce à son spectre de dichroïsme circulaire inhabituel [150]. Elle diffère très nettement des formes A et B puisqu'elle se caractérise par une double hélice gauche. Ses paires de bases sont tournées de 180° autour de l'axe (Z) par rapport à celles des formes A et B (figure 2.9). Elle présente une forte symétrie dinucléotidique accentuée par les séquences alternées RY. L'hélice a un pas

⁴Zif 268 est un facteur de transcription découvert chez la souris qui comprend trois motifs en doigt de Zinc.

⁵Le facteur TATA ou *TATA-box binding Protein* (TBP) est une des protéines formant le complexe TFIID, facteur général de la transcription qui permet la fixation de l'ARN-polymérase de type II à l'ADN.

de 45 Å et douze paires de bases par tour, elle est donc étirée par rapport aux formes A et B (28 Å/onze paires de bases et 34 Å/dix paires de bases, respectivement). Le plissement des sucres et l'angle glycosidique alternent respectivement entre C2'-endo (*anti*) et C3'-endo (*syn*) lors du passage d'une pyrimidine à la purine adjacente. Les bases sont rejetées vers l'extérieur de la double hélice du côté du grand sillon, faisant pratiquement disparaître ce dernier. Il apparaît donc large et peu profond tandis que le petit sillon est étroit et très profond (figure 2.14).

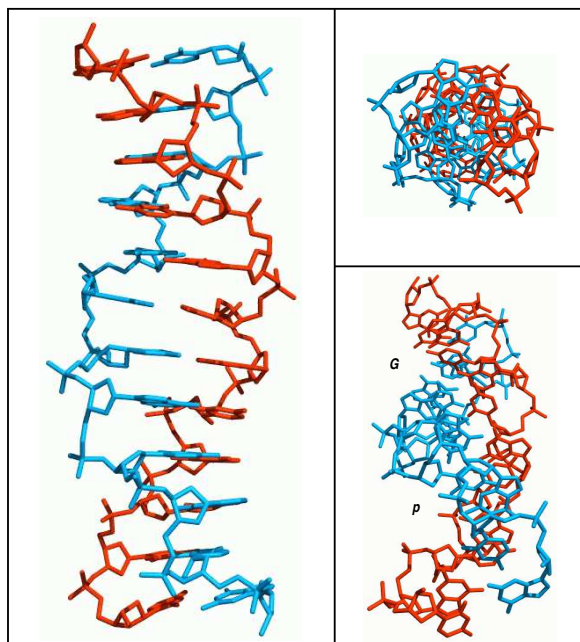


FIG. 2.14 – ADN Z

L'ADN Z est instable dans les conditions physiologiques. Il est favorisé par des séquences alternées GC, par une forte concentration saline ou par la présence d'ions multivalents [9, 99, 105, 117, 121, 195] qui écrantent les groupements phosphate et permettent donc le resserrement du petit sillon caractéristique de cette conformation. Comme les formes A et B, elle serait stabilisée par une « épine dorsale » structurée de molécules d'eau qui s'établirait dans le petit sillon [72]. Elle peut aussi être stabilisée par la modification chimique des bases comme la méthylation de la cytosine en C5 [12, 70] ou sa bromination [36, 123]. L'ADN Z peut également être induit par la création de supertours négatifs ou par réaction avec des anticorps spécifiques et certains oligopeptides [184]. Parallèlement, d'autres types de modifications chimiques, entre autres celles qui gênent la rotation des paires de bases, défavorisent la transition B→Z.

Les mécanismes de la transition B→Z sont très peu connus malgré les différents modèles proposés [78, 82, 84, 112]. Ceux-ci suggèrent qu'après une lente étape d'initiation [159], la transition se propage de paires de bases en paires de bases le long de la

séquence sans ouverture de celles-ci. La zone de jonction serait très limitée dans l'espace [53, 93, 172] et caractérisée par une importante dislocation et une forte courbure de l'axe hélicoïdal [111]. Cette rupture de l'axe serait couplée à une augmentation du *Rise* jusqu'à la perte des interactions de *stacking*, ce qui faciliterait le retournement des bases et le changement de sens de la double hélice.

2.6 Conclusion

Les données structurales dont la qualité ne cesse de s'améliorer montrent que la conformation de l'ADN est en réalité beaucoup plus complexe que la double hélice symétrique de Watson et Crick. L'ADN est une molécule relativement flexible qui présente des conformations locales diverses, dépendantes de la séquence en bases. La variabilité des conformations cristallographiques des ADN B témoigne de cette flexibilité [86], et les études par spectroscopie RMN permettent également d'obtenir des informations sur la flexibilité de l'ADN en solution [81]. Ainsi, il faut voir les trois formes canoniques de l'ADN comme trois familles de structures plutôt que comme des structures uniques. Elles signalent des zones de stabilité dans l'espace conformationnel très complexe de la double hélice. A ces familles canoniques, s'ajoute une série de structures irrégulières telles les triples ou quadruples hélices, les ADN cruciformes [81, 116] et les « fermetures éclair » d'adénines [173]. Il faut noter que toutes les structures irrégulières, ainsi que les formes A et Z sont généralement moins stables que l'ADN B dans les conditions physiologiques. Leur rôle biologique est un sujet d'étude très actif car sa compréhension est essentielle pour caractériser complètement les fonctions biologiques de l'ADN. Parmi l'ensemble de ces déformations, la courbure intrinsèque de l'ADN joue un rôle important dans les mécanismes de reconnaissance spécifique ADN - protéine.

Chapitre 3

Simulation de dynamique moléculaire

Si la mécanique moléculaire a longtemps été la panacée des représentations numériques pour les modèles atomiques, elle était limitée en pratique par les performances des ordinateurs. Depuis le début des années 90 et la progressive augmentation des performances des processeurs 32-bits économiques, celle-ci a laissé place à la dynamique moléculaire. Il est actuellement possible de réaliser des simulations de dynamique moléculaire d'une dizaine de nanosecondes, échelle de temps intermédiaire entre les conformations statiques issues des structures cristallographiques et celles dynamiques issues des structures par spectroscopie RMN. Je détaillerais dans le chapitre V les possibilités et limites de cette approche dans le cadre de la modélisation d'ADN platinés.

3.1 Méthodologie de la Dynamique Moléculaire

Le formalisme de la dynamique moléculaire utilise la seconde loi de Newton (équ. 3.1) pour décrire le mouvement d'une molécule en fonction du temps :

$$\vec{F}_i = m_i \cdot \vec{a}_i \quad (3.1)$$

- \vec{F}_i est le vecteur force agissant sur l'atome i ;
- m_i est la masse de l'atome i ;
- \vec{a}_i est le vecteur accélération de l'atome i .

Comme le montre l'équation de Newton, la vitesse et le sens du mouvement atomique dépendent des forces qui s'exercent entre les atomes. Au début d'une simulation une vitesse initiale propre est attribuée aux atomes, pour se conformer à l'énergie cinétique totale du système. Cette énergie dépend elle-même de la température de simulation désirée. Elle est habituellement obtenue en élevant lentement la température d'un système moléculaire placé initialement au zéro absolu et en amenant ainsi progressivement les atomes qui le constituent à l'énergie d'équilibre. La dynamique moléculaire permet de

calculer la force exercée sur chaque atome et de déterminer différentes informations liées à la connaissance de cette donnée (vitesse et position des atomes pendant un intervalle de temps spécifique, de l'ordre de 1 à 1000 fs suivant les systèmes moléculaires).

La force F_i qui s'exerce sur un atome i se trouvant en position $r_i(t)$ est déterminée par différenciation de la fonction énergie potentielle au voisinage de la position $r_i(t)$ (équ. 3.2) :

$$\vec{F}_i = -\frac{d\vec{E}(r_i, \dots, r_n)}{dr_i(t)} \quad (3.2)$$

L'énergie potentielle totale du système, $E(r)$, se calcule en utilisant les techniques de mécanique moléculaire ou de chimie quantique. La mécanique moléculaire ne permet de calculer l'énergie que pour des applications n'induisant pas de modifications importantes de la structure électronique (ex : rupture/formation d'une liaison). La chimie quantique donne une valeur de l'énergie permettant d'étudier des processus dynamiques entraînant des changements chimiques mais cette application est d'utilisation limitée (compte-tenu de la durée des calculs nécessaires à la détermination de la fonction énergie).

La dynamique moléculaire permet d'étudier la trajectoire des atomes à partir de la connaissance des forces s'exerçant sur chaque atome et des masses atomiques. Pour l'ensemble des atomes définissant une structure moléculaire, l'équation de Newton suivante est vérifiée (équ. 3.3) :

$$\vec{F}_i = m_i \frac{d^2 r_i(t)}{dt^2} \quad (3.3)$$

Cependant par suite d'un défaut de solution analytique, les trajectoires ne sont pas directement déterminées par résolution de l'équation de Newton précédente. Une approche différente est utilisée. Elle consiste à calculer :

- les accélérations atomiques à partir des forces s'exerçant sur chaque atome et des masses atomiques ;
- les vitesses de chaque atome à partir de la connaissance des accélérations atomiques (équ. 3.4) :

$$\vec{a}_i = \frac{d\vec{v}_i}{dt} \quad (3.4)$$

- les positions des atomes à partir des vitesses atomiques (équ. 3.5) :

$$\vec{v}_i = \frac{d\vec{r}_i}{dt} \quad (3.5)$$

L'intégration de ces équations se fait en subdivisant la trajectoire en une série d'états discrets séparés par des intervalles de temps très courts dont la longueur définit le pas d'intégration δt . Le déplacement d'un atome pendant l'intervalle de temps δt est ainsi

donné par l'équation (éq. 3.6) :

$$\vec{r}_i(t + \delta t) - \vec{r}_i(t) = \frac{d\vec{r}_i(t^*)}{dt} \Delta t = \vec{v}_i(t^*) \Delta t \quad (3.6)$$

La valeur de l'instant t^* fournie par la méthode de Verlet[193] est le milieu de δt et non $(t + \frac{\delta t}{2})$ car la vitesse n'est pas constante sur cet intervalle de temps. Il est alors possible de déterminer l'accélération à partir de la seconde loi de Newton et la vitesse propre à chaque atome (éq. 3.7) :

$$\vec{v}_i(t + \frac{\Delta t}{2}) = \vec{v}_i(t - \frac{\Delta t}{2}) + \vec{a}_i(t) \Delta t \quad (3.7)$$

La détermination de la vitesse permet d'obtenir la position de l'atome à l'aide de l'équation précédemment énoncée à l'instant $(t + \delta t)$. La répétition de cette procédure à des intervalles de temps discrets en fonction de la vitesse aboutit à l'identification de la trajectoire (tableau 3.1).

Temps	$t - \frac{\delta t}{2}$	t	$t + \frac{\delta t}{2}$	$t + \delta t$	$t + \frac{3\delta t}{2}$
Positions		$r_i(t)$		$r_i(t + \delta t)$	
Vitesses	$v_i(t - \frac{\delta t}{2})$		$v_i(t + \frac{\delta t}{2})$		$v_i(t + \frac{3\delta t}{2})$
Forces		$F_i(t)$		$F_i(t + \delta t)$	

TAB. 3.1 – Déroulement du calcul d'un pas d'intégration de dynamique moléculaire.

Notons que la cohérence des résultats obtenus par la méthode de dynamique moléculaire sera directement liée à deux éléments :

- fiabilité des potentiels ;
- valeur donnée au pas d'intégration δt : l'intervalle de temps choisi devra être plus petit que la période du mouvement de la plus haute fréquence. Son ordre de grandeur est de 1 fs, ce qui permet de simuler correctement les mouvements de vibration des liaisons hydrogènes.

3.2 Simulation en solvant explicite

Dans le cadre d'une simulation en solvant explicite (c'est-à-dire qui contient une description ponctuelle des ions et molécules d'eau), le calcul des forces devient vite prohibitif en ressources informatiques.

Pour conserver la cohérence des résultats tout en diminuant l'énorme puissance de calcul nécessaire, deux méthodes clés ont été développées.

3.2.1 Conditions périodiques et seuil de calcul

Le traitement correct des interactions à longue distance dans les simulations de dynamique moléculaire est difficile à cause du caractère à longue distance ($\frac{1}{r}$) du potentiel de Coulomb. Dans les simulations actuelles, il est ainsi nécessaire d'utiliser un « seuil de calcul » ou « *cutoff* » qui limite entre 8 et 16 Å la distance au-delà de laquelle les interactions entre un atome et ses voisins ne sont pas calculées. Toutefois, ce seuil de calcul engendrerait de graves distorsions dans le traitement correct des interactions s'il n'était pas corrigé par l'utilisation de conditions périodiques.

Considérons N atomes de charge ponctuelle q_1, q_2, \dots, q_N situés aux positions r_1, r_2, \dots, r_N dans la boîte centrale U de côté L (la boîte est supposée cubique pour la démonstration), l'ensemble étant neutre électriquement. L'énergie électrostatique totale (coulombienne) de la boîte U ainsi que les interactions entre la boîte centrale et toutes les images (conditions périodiques) est donnée par l'équation 3.8 :

$$E(r_1, r_2, \dots, r_N) = \frac{1}{2} \sum_{|n|=0}^{\infty *} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{4\pi\epsilon_r \epsilon_0 |r_{ij} + n|} \quad (3.8)$$

Le vecteur n est défini par $n = (n_x L, n_y L, n_z L)$ et le symbole « $*$ » sur la première somme indique que les termes correspondant à $i=j$ sont omis quand $n=0$. En fait, à cause du caractère longue distance du potentiel de Coulomb ($\frac{1}{r}$), calculer la somme de cette équation devient inapplicable. L'astuce pour contourner ce problème est de couper les interactions en deux termes (équation 3.9) :

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1 - f(r)}{r} \quad (3.9)$$

Le but est donc de trouver une fonction $f(r)$ appropriée, laquelle peut traiter les variations rapides de $\frac{1}{r}$ pour des distances petites (r petit) et les variations lentes de $\frac{1}{r}$ pour les longues distances (r grand). Cela veut dire que le premier terme $\frac{f(r)}{r}$ devrait être négligeable au-delà d'un certain seuil r_{max} et que le deuxième terme devrait changer lentement pour toute valeur de r permettant ainsi un calcul efficace, dans l'espace réciproque, de sa contribution au potentiel électrostatique.

La méthode dite « somme d'Ewald » [61] a été introduite en 1921 comme une technique pour calculer efficacement la somme des interactions électrostatiques entre les particules d'un système moléculaire dans les conditions périodiques.

Cette méthode calcule l'énergie électrostatique coulombienne dans des conditions périodiques en la somme de trois termes de rapide convergence :

- la somme directe qui est une somme sur l'espace réel,
- la somme réciproque (imaginaire ou de Fourier),
- la somme de correction qui est une constante de correction.

Ces trois sommes sont présentées dans les équations 3.10, 3.11, 3.12, 3.13 :

$$E_{directe}(r_1, r_2, \dots, r_N) = \frac{1}{2} \sum_{|n|=0}^{\infty} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{4\pi\epsilon_r\epsilon_0} \frac{erfc(\alpha|r_{ij} + n|)}{|r_{ij} + n|} \quad (3.10)$$

α est une constante arbitraire qui détermine le rapport relatif de convergence entre les sommes directe et réciproque.

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp(-t^2) dt \quad (3.11)$$

$$E_{reciproque}(r_1, r_2, \dots, r_N) = \frac{1}{2} \sum_{k \neq 0} \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{4\pi\epsilon_r\epsilon_0} \frac{4\pi^2}{k^2} \exp\left(-\frac{k^2}{k\alpha^2}\right) \cos(k.r_{ij}) \quad (3.12)$$

$$E_{correction} = -\frac{\alpha}{\sqrt{\pi}} \sum_{j=1}^N \frac{q_j^2}{4\pi\epsilon_r\epsilon_0} \quad (3.13)$$

L'algorithme original d'Ewald consomme un temps de calcul proportionnel à N^2 , où N est le nombre de particules du système étudié. Des approximations basées sur des seuils et l'utilisation de transformées rapides de Fourier permettent d'arriver aux temps de calcul en $N \cdot \log(N)$. La méthode PME (*Particle Mesh Ewald*) est l'un des algorithmes utilisant ces approximations [48].

Dans la méthode PME, la somme directe est évaluée explicitement en utilisant un seuil de 8 à 10 Å tandis que la somme réciproque est approximée, par l'intermédiaire des transformées rapides de Fourier (FFT : *Fast Fourier Transforms*), à une grille où les charges sont interpolées aux points de la grille en utilisant la technique d'interpolation dite « *B-spline* ». C'est l'efficacité et l'exactitude du calcul de la somme réciproque qui donne à PME son avantage informatique. Grâce à cette méthode, les simulations de dynamique moléculaire de systèmes biologiques (ADN, ARN, protéines, etc.) peuvent, depuis une décennie donner une description cohérente des résultats expérimentaux [7, 30, 31, 32, 33, 34, 166]

3.2.2 Restrictions sur la longueur des liaisons

Le mouvement de plus rapide amplitude est la vibration de la longueur des liaisons hydrogènes. Son ordre de grandeur étant de 1 fs (10^{-15} s), cette durée est donc le pas d'intégration minimum nécessaire pour une représentation correcte des forces lors d'une simulation de dynamique moléculaire. Afin d'accélérer les calculs, Berendsen et ses collaborateurs ont restreint la mobilité de ces liaisons [164], ce qui permet d'utiliser un pas d'intégration de 2fs au lieu de 1fs, divisant par deux le temps de calcul nécessaire.

Cette contrainte associée à l'utilisation de molécules d'eau de type TIP3P [100], dont le rayon de Van der Waals des hydrogènes est nul, permet de rendre ces dernières complètement rigides et assimilables au diamètre de l'oxygène seul. Cette combinaison de simplification et rigidification du modèle aqueux permet de diminuer le nombre de calculs à effectuer.

3.3 Le champ de force d'AMBER

AMBER [47, 200] utilise un modèle de représentation des molécules dit « tout atome » (« *all-atoms force field* »), toutes les caractéristiques des atomes (longueur des liaisons inter-atomiques, charge des atomes, angles de valence et de torsion, constante de force des liaisons atome-atome, constante de répulsion-attraction) sont décrites explicitement dans un ensemble que l'on appelle *champ de force*. Le calcul de l'énergie des interactions inter atomiques est effectué en utilisant l'équation 3.14.

$$\begin{aligned}
 E_{\text{total}} = & \sum_{\text{liaisons}} K_r (r - r_{eq})^2 \\
 & + \sum_{\text{angles}} K_\theta (\theta - \theta_{eq})^2 \\
 & + \sum_{\text{torsions}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \\
 & + \sum_{i < j} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{\epsilon r_{ij}} \right]
 \end{aligned} \tag{3.14}$$

L'élongation des liaisons (r) et la distorsion des angles de valence (θ) sont en général faibles, de l'ordre de 0,05 Å pour les longueurs de liaison et de quelques degrés pour les angles de valence. Les fonctions d'énergie potentielle utilisées pour décrire ces déformations sont des fonctions paraboliques dites harmoniques, suivant la loi de Hooke. Ces fonctions sont très utilisées dans tous les champs de force et sont valides si la déformation n'est pas trop importante (dans ce cas un potentiel de Morse reproduit mieux l'énergie). L'indice « eq » représente la position d'équilibre des oscillations des longueurs de liaison et des angles de valence. Les constantes K_r et K_θ sont les constantes de force dérivées de l'analyse vibrationnelle de molécules modèles (spectroscopie IR, Raman) ou des calculs de modes normaux.

L'énergie de torsion (angle dièdre) associée à la rotation (ϕ) autour d'une liaison BC définie par quatre atomes consécutifs ABCD est exprimée sous la forme d'une fonction périodique développée en série de Fourier. V_n est la constante de torsion (bar-

rière de rotation), n la périodicité de la rotation et γ est l'angle de phase. Les données cristallographiques, la RMN et la spectroscopie micro-ondes permettent d'accéder à ces trois paramètres.

L'énergie d'interaction de Van der Waals fait partie des *interactions non liées* décrivant la somme des énergie des forces attractives et répulsives. L'expression de l'énergie d'interactions de Van der Waals est exprimée en général sous la forme d'un potentiel de *Lennard-Jones* constitué des interactions dipolaires attractives entre deux atomes i et j séparés par une distance r (énergie de dispersion de London) $\frac{B_{ij}}{r_{ij}^6}$ et du terme $\frac{A_{ij}}{r_{ij}^{12}}$ qui représente les répulsions entre les mêmes atomes à cause des recouvrements des nuages électroniques à courtes distances. Les termes A_{ij} et B_{ij} sont constants pour un couple d'atomes donné.

Les interactions électrostatiques forment l'autre composante des *interactions non liées*. Elles sont calculées d'après la loi de Coulomb, où la molécule est décrite par des charges partielles centrées sur les atomes. Ces charges sont généralement déterminées par des calculs *ab initio* ou semi-empiriques. ϵ_{ij} est la constante diélectrique comptant dans l'atténuation de l'interaction électrostatique par l'environnement (solvant ou molécule elle-même); q_i et q_j sont les charges partielles des atomes concernés par l'interaction.

3.4 Mise en place d'une simulation de dynamique moléculaire

Après avoir présenté le principe de la simulation ainsi que son implémentation dans la suite logicielle AMBER, voici une description des étapes clés de la mise en place de celle-ci.

3.4.1 Préparation du composé

Quand un composé a une géométrie classique (ADN et protéine natifs), il est possible d'utiliser les outils d'AMBER comme xleap/Moil-View [26, 176] ou antechamber [197] qui permettent d'automatiser tout ou partie du processus pour créer la conformation de départ pour la dynamique.

Dans le cas du cisplatine, qui provoque des déformations structurales importantes sur l'ADN, il est nécessaire de partir de structures expérimentales pour respecter la conformation induite. Ainsi le modèle issu de l'étude aux rayons X (1AIO [181, 183] ou 1CKT [139]) a été utilisé comme point de départ.

Afin d'obtenir un jeu de coordonnées utilisables dans AMBER, il faut :

- uniformiser les types d’atomes entre la structure issue de la pdb (dont les noms des atomes du sucre est identifié par le symbole « * » : C₂* au lieu de « ' » : C₂' pour AMBER) ;
- ajouter les atomes d’hydrogènes et vérifier leur géométrie (pour éviter les inversions $H'_2 \leftrightarrow H''_2$) ;
- enlever les atomes superflus (en général hétéroatome comme le brome, utilisé en cristallographie) ;
- transformer la séquence de la structure initiale en séquence cible (en changeant les atomes dans le cadre du remplacement d’une purine par une purine ou d’une pyrimidine par une pyrimidine, soit en changeant complètement la base).

Ces opérations sont effectuées dans le fichier pdb initial (par exemple pour le remplacement des « * ») puis dans un logiciel dédié (xleap/Moil-View) qui va permettre une manipulation plus aisée des atomes, ainsi qu’une minimisation de la structure finale ainsi obtenue.

Cette structure de départ ne sert qu’à fournir les coordonnées initiales pour le modèle que nous allons chercher à minimiser, il faut ensuite :

1. intégrer les caractéristiques du complexe avec ces données expérimentales (charges, constantes de force) ;
2. ajouter les contre-ions (sodium dans le cas de l’ADN) pour neutraliser électriquement le complexe¹ ;
3. ajouter les molécules d’eau en quantité suffisante pour laisser une épaisseur aqueuse d’au moins 10 Å autour du soluté, condition sans laquelle les molécules images pourraient se retrouver en contact l’une avec l’autre, ce qui entraîne une interruption du calcul.

Le système ainsi obtenu comprend en moyenne 4500 molécules pour 15000 atomes mais dans une géométrie très régulière (qui vient du mode d’insertion des molécules des contre-ions et des molécules d’eau). Afin d’optimiser la géométrie de l’ensemble, le système est progressivement minimisé sous contraintes puis avec les contraintes relâchées (l’eau et les contre-ions sont contraints, puis seulement les molécules d’eau, puis plus de contraintes). La structure ainsi obtenue est le point de départ pour la simulation de dynamique moléculaire.

¹Il existe une petite controverse concernant l’addition de contre-ions ou non. En effet ceux-ci entraînent des perturbations durant la dynamique, qui ne sont pas généralement souhaitées, comme les variations locales de structure (les molécules d’eau s’agencent autour des contre-ions), les chocs avec le soluté. Ces inconvénients apparaissent généralement au début de la simulation et leurs effets s’atténuent avec l’allongement de celle-ci. Il est possible de ne pas mettre de contre-ions, le programme principal de simulation sander se chargeant alors de neutraliser l’ensemble soluté-solvant. Cela évite alors les effets décrits ci-dessus, mais aucune étude approfondie n’a été réalisée dans ces conditions.

3.4.2 Chauffage

L'objectif est d'amener de l'énergie à une structure statique à 0 K pour obtenir une description dynamique à 300 K² de son comportement. Il est atteint en deux étapes :

1. A 0 K, les vitesses sont calculées suivant une distribution de type Maxwell-Boltzmann [16]. Entre 0 et 150 K la simulation est effectuée à volume constant et la température est appliquée de manière stricte au système (moyennant une marge de 2 K) pour ajuster la vitesse des atomes. Le soluté et les ions de neutralisation sont contraints.
2. Entre 150 K et 300 K, les vitesses sont ajustées en fonction de la température suivant l'algorithme de Berendsen [16], le système étant maintenant sous pression constante. Le soluté et les ions de neutralisation sont contraints.

3.4.3 Equilibration

Le système étant à 300 K, il faut maintenant relâcher les contraintes sur le soluté et les contre-ions. Elles sont progressivement enlevées. Quand celles-ci ne sont plus présentes, une redistribution de type Maxwell-Boltzmann des vitesses est réalisée pour éviter le problème « hot solvent/cold solute » [83]. Une courte simulation est réalisée, puis la redistribution est réalisée. Cette dernière étape est réalisée deux fois.

Le système est maintenant à la température correcte pour la simulation, les vitesses sont maintenant correctement initialisées.

3.4.4 Production

La simulation est calculée pour une période de temps prédéterminée, suivie au cours de son avancement afin de détecter d'éventuels problèmes. Les analyses préliminaires peuvent commencer, éventuellement des contraintes peuvent être appliquées afin d'évaluer l'impact d'un changement conformationnel sur l'ensemble de la molécule étudiée.

3.4.5 Analyse

Deux types sont réalisées (dans le cas de l'ADN) :

- systématique : calcul des propriétés structurales de l'ADN (courbure, paramètres hélicoïdaux) par l'intermédiaire du logiciel CURVES [112, 113], des caractéristiques du composé platiné (programmes internes, disponibles en Annexe, page 257 et sur le site <http://www.steatch.org>) ;

²Il s'agit bien de 300 K (environ 27°C) et non pas 310 K (37°C, température plus cohérente pour la simulation de processus biologiques humains) car les propriétés globales des atomes du champ de force ont été calculées pour cette température.

- spécifique : vérification de la concordance des structures calculées avec les données expérimentales³, calculs de longueurs de liaisons, recherche de sous-conformations particulières, etc.

C'est cette longue étape qui permet de déterminer la précision du modèle obtenu ainsi que la pertinence des calculs réalisés.

³Il est toutefois toujours délicat de comparer des caractéristiques structurales n'ayant pas la même échelle de temps ni la même marge d'erreur, il faut donc se concentrer sur celles qui sont bien établies (comme la conformation d'un sucre par exemple).

Chapitre 4

Objectifs du travail

Le travail effectué précédemment au laboratoire [59] a permis d'élucider le comportement dynamique d'un adduit platiné sur la séquence 5'-GCCG*G*ATCGC-3' / 5'-GCGAT-CCGGC-3' (G* représente une guanine platinée), qui avait été étudiée par RMN mais dont plusieurs modèles avaient dû être fournis pour expliquer les différents déplacements chimiques observés [87, 107]. Cette prouesse a été accomplie grâce à l'utilisation de la simulation de dynamique moléculaire comprenant une paramétrisation nouvelle pour la conformation des sucres de l'ADN, ce qui lui a apporté une flexibilité suffisante pour atteindre les états observés en RMN.

Chélaté sur l'ADN, le cisplatine induit une forte courbure, un déroulement de l'hélice, l'ouverture du petit sillon, ainsi que le blindage de deux protons : $C_3H'_2$ est très blindé (0,6 ppm), et dans une moindre mesure, celui de la base complémentaire $C_{17}H'_2$ (0,1-0,2 ppm).

La simulation de dynamique moléculaire effectuée par le Dr. Elizondo-Riojas a ainsi permis de montrer que l'explication résidait dans le déplacement latéral de la cytosine complémentaire (C_{17}) de la première base platinée (G_{*4}) entre deux positions extrêmes.

Dans une de la position où la cytosine « pousse » la guanine vers le squelette phosphodiester, le blindage fort du proton $C_3H'_2$ par la guanine G_{*4} se produit tandis que la courbure de cycle de G_{18} entraîne le blindage faible du second proton $C_{17}H'_2$. Dans la seconde position ces blindages disparaissent car les protons sortent des cônes respectifs de blindage des guanines. La figure 4.1 donne une représentation schématique de l'adduit.

Ces observations ayant été faites *a posteriori*, il est apparu raisonnable d'observer les caractéristiques structurales supplémentaires apportées par la dynamique moléculaire. C'est ainsi que des transitions BI \rightleftharpoons BII sur l'adénine 6 ont été mises en évidence. Ces transitions sont couplées à la rupture (BI) ou à la formation (BII) d'une liaison hydrogène entre une ammine du cisplatine et l'atome N7 de l'adénine 6. Cette sous-conformation BII nouvelle a de plus pu être proposée comme utilisée dans la re-

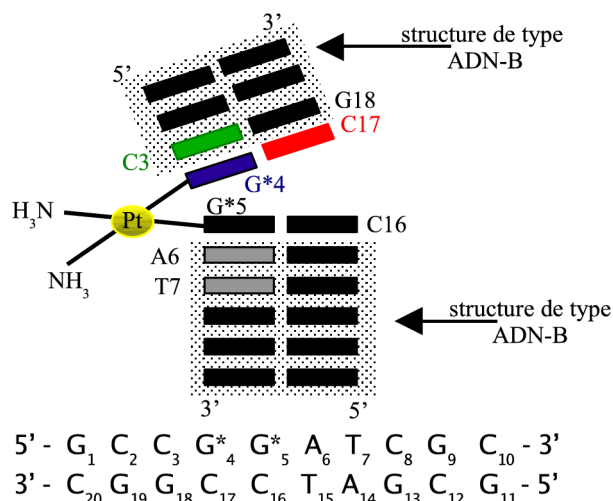


FIG. 4.1 – Représentation schématique du décanucléotide $d(GCCG^*G^*ATCGC) \cdot d(GCG \text{ ATCCGGC})$ avec l'adduit $GG\text{-}cis\text{-}Pt(NH_3)_2$ sur les guanines G^* analysé par spectroscopie RMN par Herman et ses collaborateurs [87].

connaissance entre un ADN platiné et le facteur de transcription LEF-I (à partir de la structure RMN [122]), tandis que la sous-conformation BI pourrait être reconnue par la protéine HMG-I (par identification du mode de reconnaissance issu de la structure cristallographique [139]).

Tout d'abord, pour prouver l'existence de cette sous-conformation BII, nous avons décidé de mettre au point une méthodologie permettant d'identifier et de quantifier la présence d'une conformation BII de manière plus précise et plus robuste que ce qui avait été fait jusqu'alors. Ce travail est présenté dans le premier article, chapitre I.

Ensuite, l'étude RMN de la séquence contenant la substitution de l'adénine pour la guanine a été réalisée. Cette étude est décrite dans le chapitre II. Nous y découvrirons les contraintes expérimentales qui découlent de l'utilisation de cinq guanines dans un décanucléotide, ainsi qu'une analyse des résultats mise en rapport avec la simulation de dynamique moléculaire.

Pour vérifier l'hypothèse selon laquelle la protéine LEF-I reconnaît la séquence contenant la conformation BII, une simulation de dynamique moléculaire a été réalisée en collaboration avec le Dr. Drumm. Une partie de ces travaux a donné lieu à la deuxième publication, comme présenté dans le chapitre III.

Dans le cadre de la recherche des mécanismes induisant l'efficacité des composés platinés sur les cellules tumorales, un nouvel adduit a été étudié, dans le cadre d'une collaboration avec le Dr. Teuben du laboratoire du Pr. Reedijk. Ce travail est présenté dans le troisième article, chapitre IV.

Nous verrons au travers de ces chapitres les difficultés engendrées par ces études variées et complexes puis comment l'évolution des connaissances et des technologies

nous a amené à faire infléchir notre stratégie initiale. Tout particulièrement la problématique de la représentation d'un composé chimique inorganique sera présentée dans le chapitre V.

Première partie

Détection de la conformation BII

L'étude par spectroscopie de Résonance Magnétique Nucléaire (RMN) des acides nucléiques permet de déterminer leur conformation en solution. L'enregistrement des couplages NOESY suivi de la construction d'une matrice de relaxation permet d'obtenir la structure tridimensionnelle représentative de la conformation. L'utilisation des couplages COSY et TOCSY permet de préciser les informations recueillies et de lever les ambiguïtés.

L'utilisation de ces différentes méthodes couplée à un logiciel de modélisation moléculaire permet de résoudre la structure. Cependant cette approche classique peut-être insuffisante si :

1. La structure étudiée comprend des transitions entre deux états distincts indépendants (cas $BI \rightleftharpoons BII$) ;
2. des phénomènes de diffusion de spins perturbent la qualité du signal (atténué par l'utilisation de temps de mixage longs) ;
3. le nombre de signaux de la zone d'intérêt est insuffisant pour distinguer les sous-états ;
4. un élément externe (comme la coordination du cisplatine sur l'ADN) vient perturber les signaux RMN classiques. Cela peut aider à affiner la zone d'intérêt (par exemple via l'utilisation d'azote ^{15}N qui permettra d'ajouter du signal).

Le cas des sous états BI et BII de l'ADN [67] est un exemple des difficultés rencontrées lors de la résolution de structures RMN. L'information pour identifier le sous-état présent peut-être obtenue par l'enregistrement 1D du spectre du phosphore et par couplage scalaire $^3J_{H'_3-P}$.

Cependant, comme montré par Goreinstein et son équipe [75], l'ambiguïté entre les deux sous-conformations ne peut être levée en appliquant l'équation reliant le couplage dipolaire à la conformation [102]. Ces difficultés nous ont amené à développer une approche complémentaire pour la détermination de la présence des sous-conformations BII.

Cette approche est basée sur l'analyse statistique de structures issues d'instantanés minimisés de simulations de dynamiques moléculaires et de structures cristallographiques à haute résolution ($\leq 2,0$ Å). Le nombre de pas étudiés dans le cadre de ce travail est de 26978 pour les structures issues des simulations et de 548 pas pour les structures cristallographiques.

Nous avons analysé les caractéristiques différenciant les pas BI des pas BII en terme de distances inter-protons. Cette approche a déjà été utilisée antérieurement [187] et a été étendue dans le cadre de cette étude pour prendre en compte les 21 distances possibles dans un pas $X_n - X_{n+1}$ (où X représente une base parmi A, T, C, G). Les résultats obtenus sont présentés dans le tableau 1.

	Dynamique moléculaire		Cristallographie	
Distance	B1	B2	B1	B2
H1' (n)-H6/H8 (n)	3.82 (08)	3.81 (07)	3.80 (09)	3.77 (08)
H2' (n)-H6/H8 (n)	2.40 (23)	2.30 (09)	2.37 (29)	2.15 (15)
H2'' (n)-H6/H8 (n)	3.78 (18)	3.18 (16)	3.77 (25)	3.16 (16)
H3' (n)-H6/H8 (n)	4.18 (31)	4.65 (10)	4.01 (37)	4.49 (21)
H4' (n)-H6/H8 (n)	4.76 (17)	5.34 (13)	4.73 (23)	5.24 (17)
H5' (n)-H6/H8 (n)	4.26 (35)	5.24 (24)	4.25 (32)	4.95 (27)
H5'' (n)-H6/H8 (n)	4.99 (32)	5.91 (21)	4.89 (31)	5.60 (28)
H1' (n)-H6/H8 (n+1)	3.95 (41)	2.82 (17)	3.83 (58)	3.42 (57)
H2' (n)-H6/H8 (n+1)	3.41 (36)	4.59 (25)	3.31 (44)	4.80 (44)
H2'' (n)-H6/H8 (n+1)	2.48 (31)	2.99 (32)	2.38 (33)	3.14 (47)
H3' (n)-H6/H8 (n+1)	4.95 (31)	4.38 (14)	4.70 (33)	4.52 (18)
H4' (n)-H6/H8 (n+1)	6.16 (32)	4.44 (19)	5.97 (32)	5.03 (35)
H5' (n)-H6/H8 (n+1)	7.38 (23)	6.64 (13)	7.30 (19)	7.11 (34)
H5'' (n)-H6/H8 (n+1)	7.25 (25)	6.24 (12)	7.05 (25)	6.65 (23)
H1' (n+1)-H6/H8 (n+1)	3.80 (08)	3.88 (06)	3.79 (09)	3.82 (09)
H2' (n+1)-H6/H8 (n+1)	2.39 (24)	2.36 (12)	2.35 (27)	2.29 (42)
H2'' (n+1)-H6/H8 (n+1)	3.67 (30)	3.78 (11)	3.68 (32)	3.66 (30)
H3' (n+1)-H6/H8 (n+1)	4.25 (38)	4.30 (18)	4.09 (41)	4.04 (41)
H4' (n+1)-H6/H8 (n+1)	4.86 (30)	4.80 (13)	4.80 (30)	4.82 (29)
H5' (n+1)-H6/H8 (n+1)	4.54 (44)	3.80 (27)	4.37 (43)	4.11 (26)
H5'' (n+1)-H6/H8 (n+1)	5.22 (45)	4.72 (25)	5.02 (42)	4.78 (25)
Nombre de pas	22212	4766	451	91

TAB. 1 – Récapitulatif des 21 distances mesurées sur les structures issues des données cristallographiques et des simulations de dynamique moléculaire. Les distances indiquées en rouge sont celles présentant le plus d'intérêt dans le cadre de notre étude.

Il est intéressant de noter la très bonne correspondance entre les distances observées en dynamique moléculaire et celles observées par cristallographie aux rayons X (coefficient de détermination $R=0,99$ sur la moyenne pour BI, $R=0,95$ sur la moyenne pour BII).

De ces 21 distances, seules 4 peuvent être retenues pour la séparation des deux sous-états, elles comportent à la fois une différence significative entre les deux états et sont suffisamment petites pour être identifiables correctement en RMN (si la distance se rapproche de 5 Å, l'effet Overhauser ne permet pas d'obtenir une précision suffisante sur la distance). Il s'agit de :

- H2'' (n)-H6/H8 (n)
- H1' (n)-H6/H8 (n+1)
- H2' (n)-H6/H8 (n+1)
- H2'' (n)-H6/H8 (n+1)

Les distances $H2''(n)-H6/H8(n)$, $H2'(n)-H6/H8(n+1)$ et $H2''(n)-H6/H8(n+1)$ sont sensibles à la conformation du sucre. Il est en particulier possible de distinguer les différentes populations de sucres, ainsi que de déterminer indirectement la conformation du phosphate sur le brin complémentaire.

La figure 2 présente la méthodologie de détermination de la proportion des populations BI et BII au sein d'un pas $X_n - X_{n+1}$.

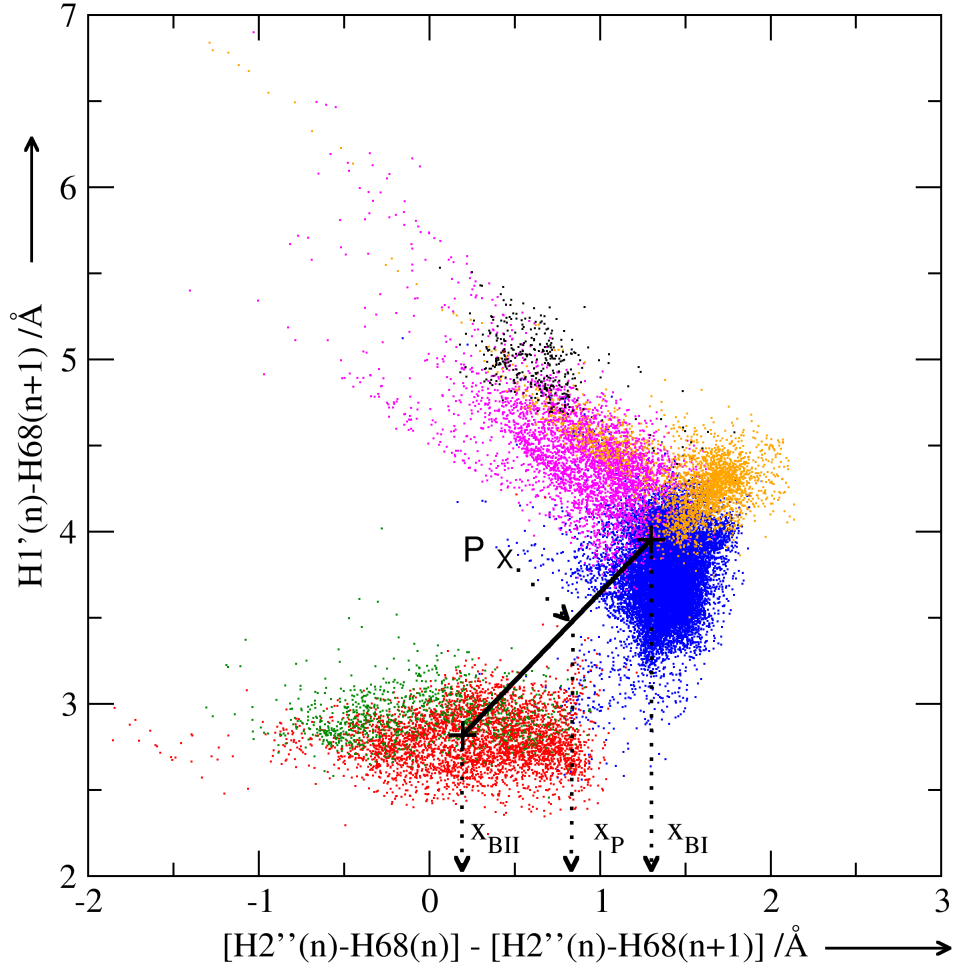


FIG. 2 – projection bi-dimensionnelle de la différence entre les distances $H2''(n)-H6/H8(n)$ et $H2''(n)-H6/H8(n+1)$ contre la distance $H1'(n)-H6/H8(n+1)$. Chaque point représente une des 26978 conformations dinucléotiques issues des simulations de dynamique moléculaire. Les couleurs indiquent les sous-conformations des pas. En bleu : BI, sucre 5' S, phosphate 3' BI. En Orange : BI, sucre 5' E, phosphate 3' BI. En noir : BI, sucre 5' N, phosphate 3' BI. En magenta : BI, phosphate 3' BII. En rouge : BII, phosphate du brin complémentaire en BI. En vert : BII, phosphate du brin complémentaire en BII. Pour déterminer le plissement du sucre, les critères suivants ont été appliqués pour l'angle de la phase P : N, $P < 60^\circ$; E, $60^\circ \leq P \leq 115^\circ$; S, $115^\circ \leq P \leq 220^\circ$. Le signe '+' indique le barycentre de chaque conformation, X_{BI} , X_{BII} la projection de ceux-ci. Le point P présente à titre d'exemple la projection obtenue pour le pas P donné.

Nous utilisons la différence entre $H2''(n)-H6/H8(n)$ et $H2''(n)-H6/H8(n+1)$ pour bénéficier du comportement antagoniste de ces distances en fonction de la conformation BI ou BII observée.

La proportion de BII est obtenue en utilisant l'équation 1 :

$$(NOE) = 100 * \frac{(X_{BI} - X_p)}{(X_{BI} - X_{BII})} \quad (1)$$

X_{BI} , X_{BII} : projection des barycentres calculés à partir des simulations de dynamique moléculaire de BI et BII sur l'axe x.

X_p : projection des distances expérimentales sur l'axe x pour le pas P.

Cette méthode de calcul a été utilisée pour prédire la proportion de BII et comparer ce résultat avec la proportion réelle. Les points de comparaison ont été extraits de la simulation de dynamique moléculaire et d'un ensemble filtré de structures cristallographiques à haute résolution. La marge d'erreur de la méthode a ainsi pu être estimée à 20 %.

A travers cette étude, une nouvelle approche a été développée pour déterminer de manière indirecte la proportion de la sous-conformation BII. Cette proportion est souvent sous-estimée en raison des contraintes expérimentales (ambiguïté dans l'identification du phosphate, couplage dipolaire altéré), mais aussi lors du raffinement des modèles (les logiciels de raffinement des molécules prennent par défaut la conformation BI).

Cette approche robuste et simple permet de déterminer si une proportion significative de la configuration BII est présente dans la structure étudiée. Cette information permettra de lever quelques ambiguïtés dans la résolution des structures, et permettra la proposition des deux modèles alternatifs avec une bonne indication quant aux proportions de chaque état.

Il reste à étendre cette étude à un ensemble plus important de structures puis à l'intégrer dans les logiciels de raffinement de RMN.

For submission to *J. Biomol. Struct. Dyn.*

**Discrimination between BI and BII conformational substates of
B-DNA based on sugar-base interproton distances**

Stéphane Teletchéa,¹ Brigitte Hartmann,² and Jiř Kozelka^{1*}

¹Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques, UMR 8601 CNRS,
Université René Descartes, 45, rue des Saints-Pères, 75270 Paris, France

²Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, UPR CNRS
9080, 13, rue Pierre et Marie Curie, 75005 Paris, France

* Corresponding author. Tel: +331 42 86 20 86; Fax: +331 42 86 20 86; E-mail: kozelka@biomedicale.univ-paris5.fr

Abstract

Molecular dynamics (MD) simulations of four water-solvated DNA duplexes were used to generate a database of ~27000 dinucleotide conformations. Analyzing this database, we investigated the relationship between so-called BI-BII transitions and short-range interproton distances. Four H-H distances were found particularly sensitive to BI-BII transitions: internucleotide H1'(n)-H68(n+1), H2'(n)-H68(n+1), and H2''(n)-H68(n+1), and intranucleotide H2''(n)-H68(n). Determination of these distances using classical NOESY spectroscopy can thus provide valuable indications on the existence of BII substates, complementing the existing method based on ^{31}P chemical shifts and ^{31}P - ^1H spin-spin-coupling constants.

Introduction

DNA flexibility is an important element of DNA-drug and DNA-protein interactions. Protein recognition of DNA usually involves transitions of both partners from their lowest-energy conformations to higher "substates". One substate frequently implicated in protein-DNA recognition is the so-called BII conformation of DNA (1-6). It differs from the canonical BI conformation in the two torsion angles ϵ and ζ (see Figure 1) which are *trans/gauche*⁻ in BI and *gauche*⁻/*trans* in BII. First reports of BII conformations in DNA came from crystallographic studies (7). Several later NMR studies provided evidence for the existence of BI \rightleftharpoons BII equilibria in solution (4,8-10). Recognition proteins alter BI \rightleftharpoons BII equilibria, and the structural aspects of this recognition are subject of current research (1,6,11,12). So far, detection of BII substates in solution has relied upon measurements of $^3\text{J}_{\text{H}3'-\text{P}}$ spin-spin coupling constants and/or ^{31}P chemical shifts, δP (9). $^3\text{J}_{\text{H}3'-\text{P}}$ is related to the torsion angle ϵ via the "modified Karplus equation" (13), while δP depends on ζ , undergoing a downfield shift when ζ passes from *gauche*⁻ (BI) to *trans* (BII) (14). However, not always can the P and H3' resonances be assigned; in addition, there are ambiguities in relating δP and $^3\text{J}_{\text{H}3'-\text{P}}$ to BI or BII (15). Therefore, additional indications, allowing an $\text{N}^n\text{pN}^{n+1}$ step between the nucleotides N^n (5'-side) and N^{n+1} (3'-side) to be assigned to BI or BII (or to a mixture of both), would be desirable.

Fedoroff et al (16) and Hartmann et al (17,18) investigated previously the dependence of proton internucleotide distances on B-DNA conformation and found strong correlation between some of the H-H distances and helicoidal parameters. Since the helicoidal parameters are, in turn,

sensitive to BI→BII backbone transitions, it should be possible to find a direct link between H-H distances and the torsion angles ϵ and ζ . In a search for such a link, we have analyzed in the present work 1766 energy-minimized structures obtained from molecular dynamics (MD) simulations of four oligonucleotides explicitly solvated in water. We present here the results of this analysis which suggest that four short-range interproton distances are particularly sensitive to BI→BII transitions. Thus, measurement of these distances by NOESY techniques is suggested to be a valuable complement to ^{31}P NMR spectroscopy for the detection of BII substates.

Methods

The MD simulations (7.5-10 ns long) were performed on four DNA duplexes (Na^+ salts) solvated in 3600-3700 TIP3P H_2O molecules: d(TCCGGGTTCT)₂, d(TCCGGATTCT)₂, d(GACTTTCAGGGT)₂, and d(GCTGGGGACTTT)₂. The calculations used the AMBER6.0 package (19) with the parm98 force-field (20). All details of the simulations were as described previously (21). Snapshots were taken every 20 ps and energy-minimized using the Generalized Born solvation model (22) with an rms gradient of $0.05 \text{ kcal.mol}^{-1}\text{\AA}^{-1}$ as convergence criterion. The 39 crystal structures with resolution of at least 2.0 \AA selected from the NDB are listed in Table S1. Hydrogen atoms were added with the PROTONATE module of AMBER and their position were refined by a belly energy minimization. The MD and X-ray structures were analyzed using CURVES (23).

Results and Discussion

The sequences of the four oligonucleotides whose MD simulations were used to generate our structural database contained 14 of the 16 possible dinucleotide steps (all except GpC and TpA). In the 1766 energy-minimized structures, each dinucleotide step (except for the terminal steps which were discarded) was classified as BI ($\epsilon\text{-}\zeta \leq 0$) or BII ($\epsilon\text{-}\zeta > 0$). This provided us with a database of 22212 BI and 4766 BII dinucleotide steps.

For all the dinucleotides of this database, we determined the $\text{Hx}(n)\text{-H68}(n+1)$ internucleotide sugar-base distances observable by NMR (Hx stands for the sugar protons H1', H2', H2'', and H3'; H68 indicates the base proton H6 or H8). In addition, we considered the intranucleotide interproton Hx-H68 distances for both nucleotides n and $n+1$. The average

distances calculated for both the BI and the BII ensembles of dinucleotide steps are shown in Table 1. We observe that all the four internucleotide distances change significantly between BI and BII. This is related to the structural changes within the 5'-nucleotide (i.e., changes in sugar pucker and the glycosidic torsion angle χ) and to the increase of twist in the BII substate (15,17,24). Whereas $H3'(n)-H68(n+1)$ is at the detection limit, the other three distances can be measured using NOESY techniques. Of the intranucleotide distances, $H2''(n)-H68(n)$ shows a significant change. Thus, in a window between 2.5 and 4.5 Å readily accessible to NOESY measurements, four interproton distances can be used to discriminate between BI and BII substates: $H1'(n)-H68(n+1)$, $H2'(n)-H68(n+1)$, $H2''(n)-H68(n+1)$, and $H2''(n)-H68(n)$ (Figure 1A).

Because of fluctuations of the four distances (see the standard deviations in Table 1), and because of the intrinsic inaccuracy of the interproton distance measurement (25,26), one distance will generally not suffice for a safe distinction between BI and BII. However, Table 1 suggests that it is possible to consider two or more distances simultaneously and increase thus the resolution between BI and BII beyond the fluctuations. This is illustrated in Figures 2A-C, where the $H1'(n)-H68(n+1)$ distance is plotted against each of the three other distances. There is virtually no overlap between BI and BII domains. Thus, the distinction between predominantly BI and predominantly BII conformations should be straightforward, provided that the distance $H1'(n)-H68(n+1)$ and one of the three other distances can be determined. In the case that $H1'(n)-H68(n+1)$ cannot be measured, the plot of the distances $H2''(n)-H68(n)$ and $H2'(n)-H68(n+1)$ against each other (Figure 2D) can be used. Figure 2 also reveals the effect of groups adjacent to the phosphate in consideration. For instance, an N sugar pucker 5' to a BI phosphate manifests itself in extremely large $H1'(n)-H68(n+1)$, $H2''(n)-H68(n+1)$, and $H2''(n)-H68(n)$ distances, and a very short $H2'(n)-H68(n+1)$ distance.

An even better separation of the BI and BII domains than that in Figure 2 was achieved by exploiting the "antagonism" between the distances $H2''(n)-H68(n)$ and $H2''(n)-H68(n+1)$. In BI, $H2''(n)-H68(n)$ is considerably larger than $H2''(n)-H68(n+1)$, whereas in BII, the intranucleotide distance becomes shorter and the internucleotide distance longer, making both distances comparable (Table 1). Figure 3 shows a plot of the difference between the $H2''(n)-H68(n)$ and $H2''(n)-H68(n+1)$ distances against the distance $H1'(n)-H68(n+1)$; the BI and BII domains are obviously still better separated than in Figure 2. An additional advantage of Figure 3 is that the

difference between two distances is generally more precisely determined than the individual distances.

The structures that we analyzed and represented in Figures 2 and 3 are snapshots from MD simulations that have been energy-minimized. In real aqueous solutions, thermal movement provides enough kinetic energy to cross the activation barrier between BI and BII, and phosphates show generally $BI \rightleftharpoons BII$ equilibria. In order to examine whether Figure 3 could be used to predict the composition of such an equilibrium, we calculated the centers of gravity for both the BI and BII ensembles in Figure 3 and joined them by a straight line, as shown. For all the non-terminal dinucleotide steps present in the four oligonucleotides (64 steps), we simulated the NOE-determined time average for each of the three distances as $\langle r^{-6} \rangle^{-1/6}$ over all energy-minimized snapshots. With these "experimental" distances, we then calculated for each of the 64 steps the difference between $H2''(n)-H68(n)$ and $H2''(n)-H68(n+1)$, and plotted it against $H1'(n)-H68(n+1)$ into Figure 3. This yielded one point for each of the 64 steps; an arbitrary example is labeled "P" in Figure 3. The obtained point was projected on the straight line and the x coordinate of the projection, x_P , was used, together with the x-coordinates of the gravicenters, x_{BI} and x_{BII} , to calculate the percentage of BII according to Eq. (1).

$$\%BII(NO E) = 100 * (x_{BI} - x_P) / (x_{BI} - x_{BII}) \quad (1)$$

If the projected point lay outside the segment delimited by the two centers of gravity, %BII(NO E) was considered to be 0 (if $x_P > x_{BI}$) or 100 (if $x_P < x_{BII}$), respectively. The calculated %BII(NO E) value is plotted in Figure 4 against the fraction of snapshots having the phosphate in consideration in the BII state, %BII(MD). In this plot, %BII(NO E) simulates the fraction that the spectroscopist would determine from NOE integration, and %BII(MD) the actual BII fraction. It can be seen that the correlation between the two values is fairly good, which suggests that measuring the three distances $H2''(n)-H68(n)$, $H1'(n)-H68(n+1)$, and $H2''(n)-H68(n+1)$ should suffice to detect not only pure BII steps but even steps with a partial population of BII.

The structures analyzed in this work were refined using the forcefield parm98 of AMBER (20). They could therefore be, in principle, biased by this forcefield. As a check, we analyzed the X-ray structures of oligonucleotides available from the Nucleic Acids Data Bank (NDB) having at least 2.0 Å resolution (39 structures; discarding terminal steps yielded a database of 548 dinucleotide steps which contained all the 16 possible dinucleotide sequences). Figure 5 displays

the plot analogous to that shown in Figure 3 for these crystal structures. The regions occupied by the different conformational domains are similar to those seen in Figure 3. The exception is the BII.BII conformation (i.e., BII with the face-to-face phosphate being also BII; green points) which in the crystal structures has a subdomain characterized by significantly longer $H1'(n)$ -H68(n+1) and $H2''(n)$ -H68(n+1) internucleotide distances. A detailed analysis revealed that this subdomain contains exclusively CpA/TpG sequences at penultimate positions, whose very particular structure has been noted previously (27,28). Whether these unusual conformations reflect specific sequence effects, crystal packing or influence of counter-ions, is not clear at this moment. Overall, the crystal structures confirm the results that we obtained from the MD simulations.

We wish to caution that the present analysis involves specifically DNA oligonucleotides belonging to the B-DNA family. Strong deviations from B-DNA type structure, e.g., towards A-DNA, can significantly alter proton-proton distances, as exemplified for the $H1'(n)$ -H68(n+1) distance in Table 2 of Ref. (29). In such cases, alteration in interproton distances arising from changes in the ϵ and ζ torsion angles will be overlaid by those originating in changes of the gross helical structure.

In Table 1, Figure 3, and Eq. 1, we have delineated simple rules allowing BII substates to be detected on the basis of routine NOESY measurements of intra- and internucleotide H-H distances. We believe that these rules provide a valuable complement to the current methods distinguishing BI from BII on the basis of δp and $^3J_{H3'-P}$. Our analysis was performed using a large number of different dinucleotide steps in different environments, and the derived rules are therefore sequence-unspecific. A refined analysis taking sequence effects into account is under progress.

Supplementary Material

Table S1: Listing the sequences and references of the 39 crystal structures selected from the NDB for analysis of dinucleotide steps are available at JBSD Online.

Acknowledgements

We thank Dragana Djuranović for the search of the PDB database. Financial support from the Association for International Cancer Research (award N° 00-321) and computer time from the IDRIS computer center of the CNRS are gratefully acknowledged. Support from COST (Action D20/003/01) has enabled us to have fruitful exchanges with other european laboratories.

References

1. Botuyan, M. V., Keire, D. A., Kroen, C. and Gorenstein, D. G., *Biochemistry* 32, 6863-6874 (1993).
2. Gorenstein, D. G., *Chem. Rev.* 94, 1315-1358 (1994).
3. Pichler, A., Rüdissler, S., Mitterböck, M., Huber, C. G., Winger, R. H., Liedl, K. R., Hallbrucker, A. and Mayer, E., *Biophys. J.* 77, 398-409 (1999).
4. Tisné, C., Delepierre, M. and Hartmann, B., *J. Mol. Biol.* 293, 139-150 (1999).
5. Isaacs, R. J. and Spielmann, H. P., *J. Mol. Biol.* 311, 149-160 (2001).
6. Hartmann, B., Sullivan, M. R. and Harris, L. F., *Biopolymers* 65, in press (2003).
7. Fratini, A. V., Kopka, M. L., Drew, H. R. and Dickerson, R. E., *J. Biol. Chem.* 257, 14686-14707 (1982).
8. Gorenstein, D., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V. and Jones, C. R., *Biochemistry* 27, 7223-7237 (1988).
9. Schroeder, S. A., Roongta, V., Fu, J. M., Jones, C. R. and Gorenstein, D., *Biochemistry* 28, 8292-8303 (1989).
10. Lefebvre, A., Mauffret, O., Lescot, E., Hartmann, B. and Femandjian, S., *Biochemistry* 35, 12560-12569 (1996).
11. Castagné, C., Murphy, E. C., Gronenborn, A. M. and Delepierre, M., *Eur. J. Biochem.* 267, 1223-1229 (2000).
12. Djuranovic, D. and Hartmann, B., *J. Biomol. Struct. Dyn.*, in press. (2003).
13. Lankhorst, P. P., Haasnoot, C. A. G., Erkelens, C. and Altona, C., *J. Biomol. Struct. Dyn.* 1, 1387-1405 (1984).
14. Gorenstein, D. G. (1984) *Phosphorus-31 NMR: Principles and Applications*, Academic Press, New York
15. Hartmann, B., Piazzola, D. and Lavery, R., *Nucleic Acids Res.* 21, 561-568 (1993).

16. Fedoroff, O. Y., Reid, B. R. and Chuprina, V. P., *J. Mol. Biol.* 245, 325-330 (1994).
17. Lefebvre, A., Fermandjian, S. and Hartmann, B., *Nucleic Acids Res.* 25, 3855-3862 (1997).
18. Tisné, C., Hantz, E., Hartmann, B. and Delepierre, M., *J. Mol. Biol.* 279, 127-142 (1998).
19. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham III, T. E., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Jr., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Tsui, V., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K. and Kollman, P. A. (1999), University of California, San Francisco
20. Cheatham III, T. E., Cieplak, P. and Kollman, P. A., *J. Biomol. Struct. Dyn.* 16, 845-862 (1999).
21. Elizondo-Riojas, M.-A. and Kozelka, J., *J. Mol. Biol.* 314, 1227-1243 (2001).
22. Tsui, V. and Case, D., *J. Am. Chem. Soc.* 122, 2489-2498 (2000).
23. Lavery, R. and Sklenar, H., *J. Biomol. Struct. Dyn.* 6, 655-667 (1989).
24. Lipanov, A., Kopka, M. L., Grzeskowiak, M. K., Quintana, J. and Dickerson, R. E., *Biochemistry* 32, 1373-1389 (1993).
25. van de Ven, F. J. M. and Hilbers, C. W., *Eur. J. Biochem.* 178, 1-38 (1988).
26. Wijmenga, S. S. and van Buuren, B. N. M., *Prog. Nucl. Mag. Res. Spect.* 32, 287-387 (1998).
27. Heinemann, U. and Alings, C., *J. Mol. Biol.* 210, 369-381 (1989).
28. Chiu, T. K. and Dickerson, R. E., *J. Mol. Biol.* 301, 915-945 (2000).
29. Trantirek, L., Stefl, R., Vorlickova, M., Koca, J., Sklenar, V. and Kypr, J., *J. Mol. Biol.* 297, 907-922 (2000).

Figure Legends

Figure 1. A: Atom and torsion angle labeling for a DNA dinucleotide. The four interproton distances $H1'(n)-H68(n+1)$, $H2'(n)-H68(n+1)$, $H2''(n)-H68(n+1)$, and $H2''(n)-H68(n)$ are marked with dashed lines. **B:** Drawings of a DNA dinucleotide in BI and BII conformations, with the hydrogen atoms omitted for clarity.

Figure 2A-D. Two-dimensional plots of interproton distances occurring in the ensemble of 26978 dinucleotide steps from four MD simulations. Each point represents one dinucleotide step and the color code indicates its conformation. Blue: **BI**, 5'-sugar S, 3'-phosphate BI. Orange: **BI**, 5'-sugar E, 3'-phosphate BI. Black: **BI**, 5'-sugar N, 3'-phosphate BI. Magenta: **BI**, 3'-phosphate BII. Red: **BII**, face-to-face phosphate BI. Green: **BII**, face-to-face phosphate BII. The sugar pucker assignment was based on the phase angle P : N, $P < 60^\circ$; E, $60^\circ \leq P < 115^\circ$; S, $115^\circ \leq P < 220^\circ$.

Figure 3. Two-dimensional plot of the difference between distances $H2''(n)-H68(n)$ and $H2''(n)-H68(n+1)$ against the distance $H1'(n)-H68(n+1)$. The ensemble of dinucleotide steps analyzed and the color code are the same as in Figure 2. "+" indicate the centers of gravity of the BI and BII domains. P represents an example of a point that would be determined from the NOE-based measurement of time-averaged distances for a given dinucleotide step. Its projection on the straight line connecting the two centers of gravity yields x_P , which is used in Eq. (1) to calculate %BII.

Figure 4. Correlation of the actual BII fractions %BII(MD) occurring in the 64 dinucleotide steps with the "NOE-based" BII fractions %BII(NOE), calculated on the basis of the r^6 -averaged distances $H2''(n)-H68(n)$, $H1'(n)-H68(n+1)$, and $H2''(n)-H68(n+1)$, using Figure 3 and Eq. (1) as described in the text. The dashed line corresponds to the equality %BII(NOE) = %BII(MD).

Figure 5. Two-dimensional plot analogous to that in Figure 3, for the ensemble of non-terminal dinucleotide steps of the 39 crystal structures selected from the NDB. See Figure 2 for the color code.

Table 1. Interproton distances averaged over the BI and BII ensembles of dinucleotide steps occurring in the 1766 energy minimized snapshots from MD simulations of 4 oligonucleotides.

	Distance / Å	
	BI	BII
H1' (n)-H6/H8 (n)	3.82 (08)	3.81 (07)
H2' (n)-H6/H8 (n)	2.40 (23)	2.30 (09)
H2'' (n)-H6/H8 (n)	3.78 (18)	3.18 (16)
H3' (n)-H6/H8 (n)	4.18 (31)	4.65 (10)
H1' (n)-H6/H8 (n+1)	3.95 (41)	2.82 (17)
H2' (n)-H6/H8 (n+1)	3.41 (36)	4.59 (25)
H2'' (n)-H6/H8 (n+1)	2.48 (31)	2.99 (32)
H3' (n)-H6/H8 (n+1)	4.95 (31)	4.38 (14)
H1' (n+1)-H6/H8 (n+1)	3.80 (08)	3.88 (06)
H2' (n+1)-H6/H8 (n+1)	2.39 (24)	2.36 (12)
H2'' (n+1)-H6/H8 (n+1)	3.67 (30)	3.78 (11)
H3' (n+1)-H6/H8 (n+1)	4.25 (38)	4.30 (18)
Number of steps	22212	4766

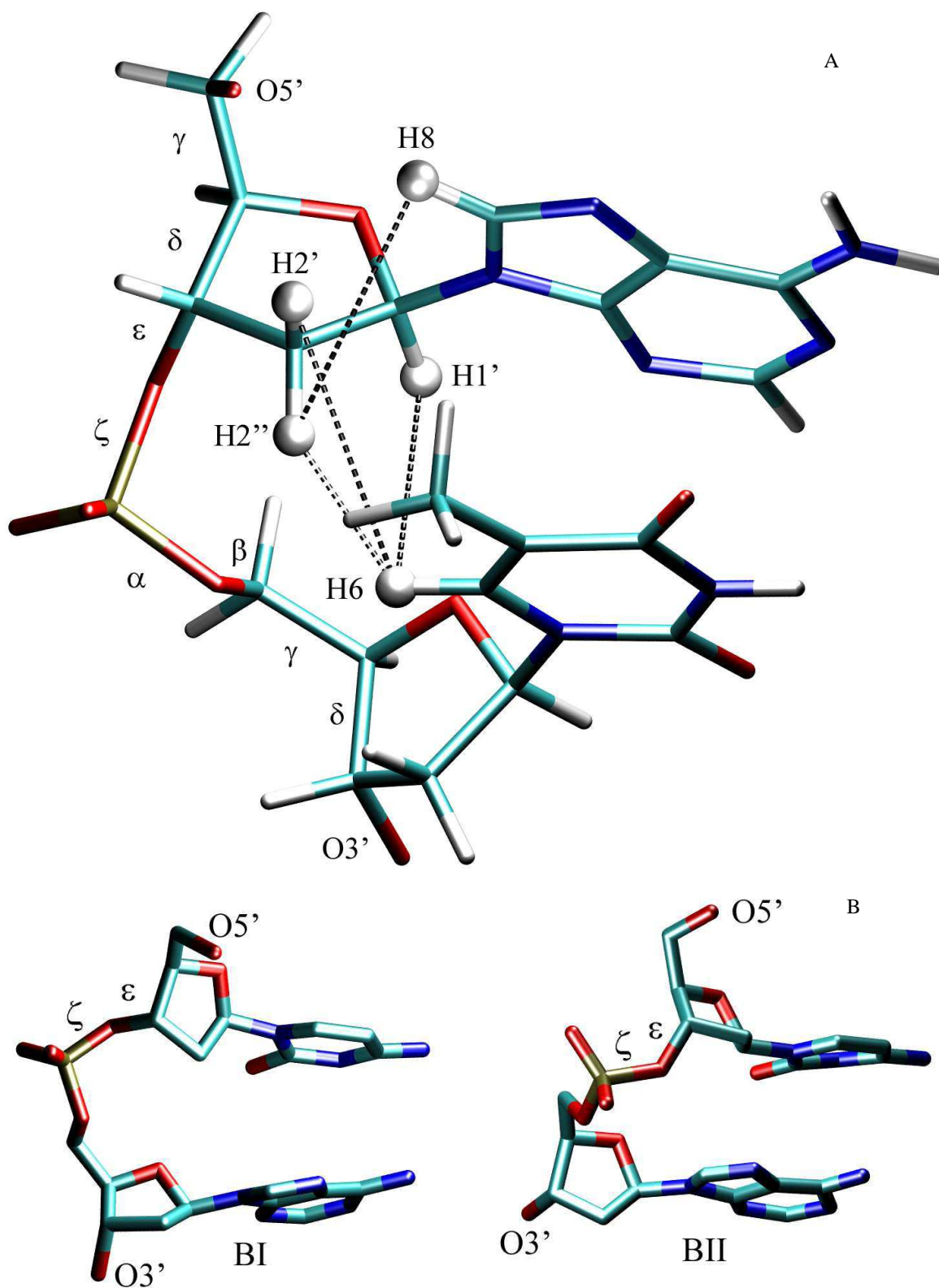
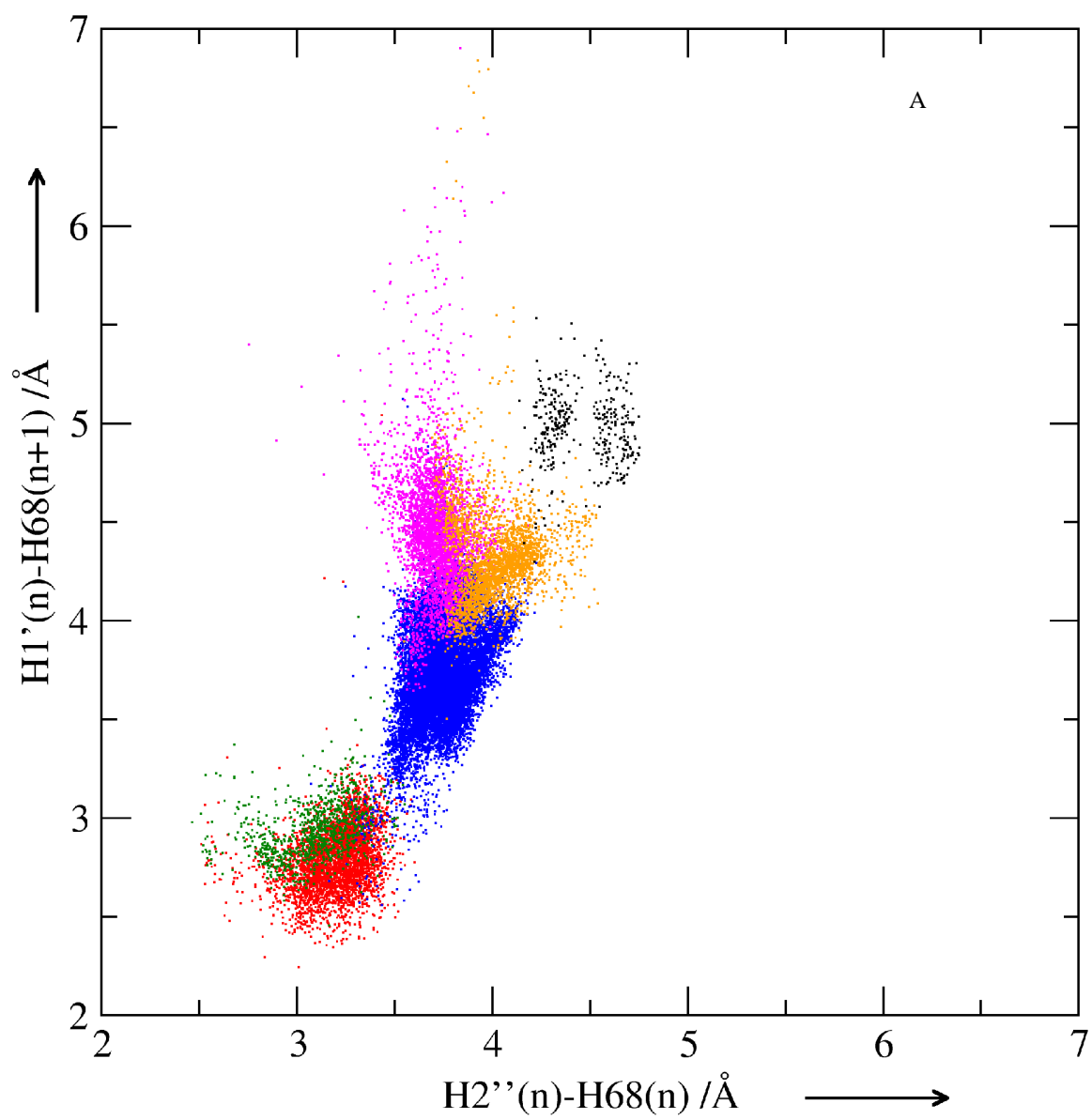
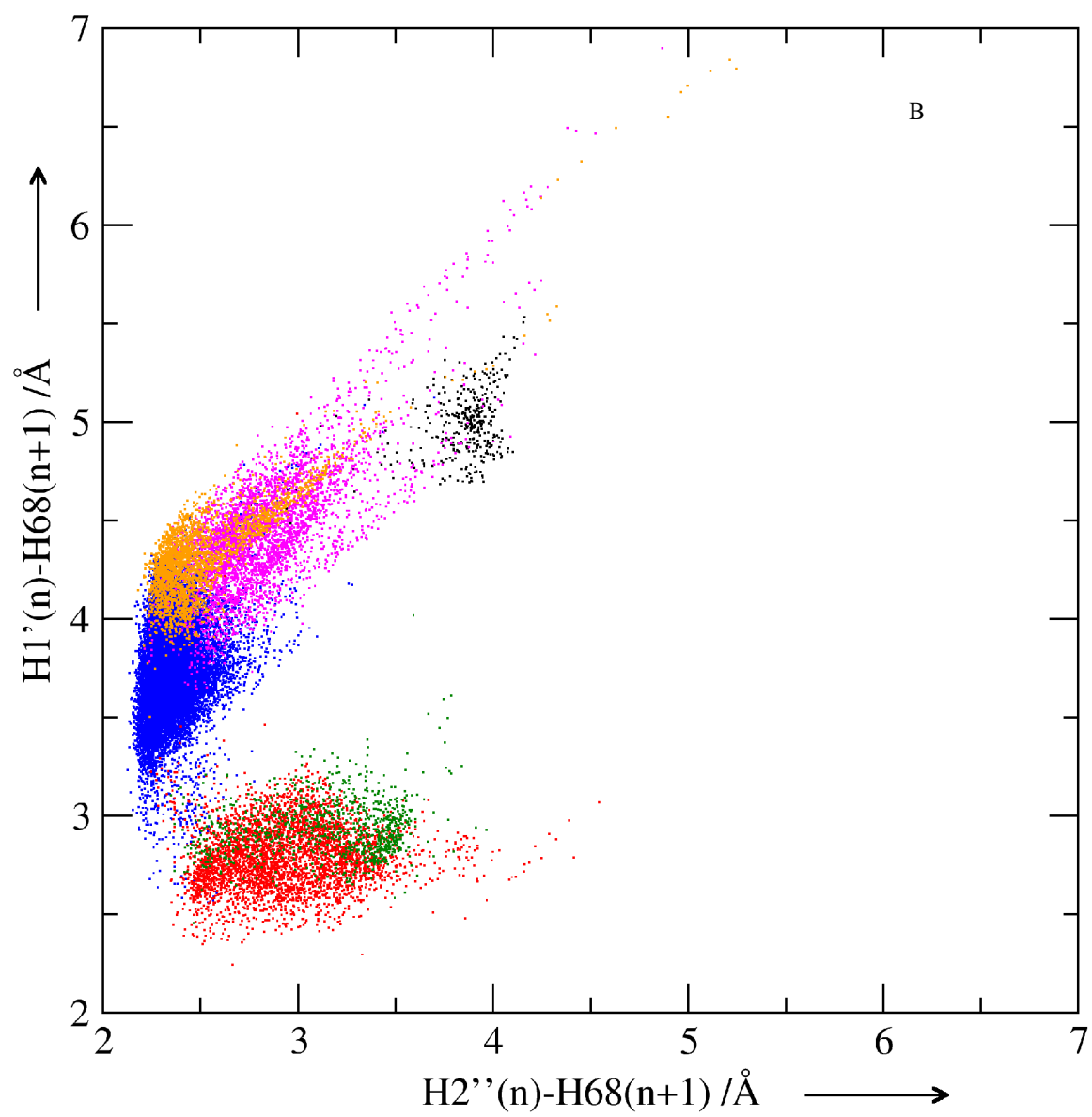
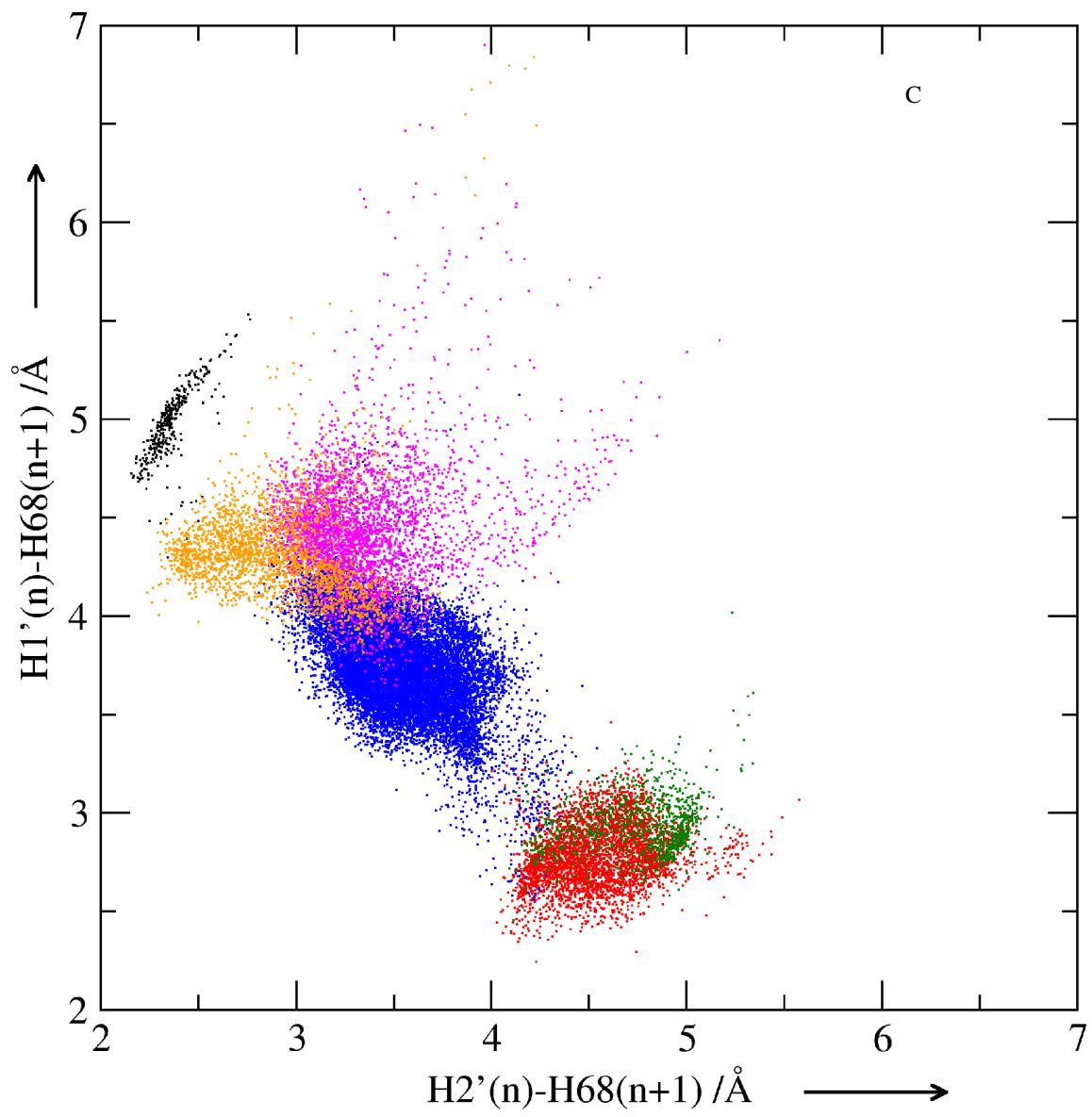


Figure 1. A: Atom and torsion angle labeling for a DNA dinucleotide. The four interproton distances $H1'(n)-H68(n+1)$, $H2'(n)-H68(n+1)$, $H2''(n)-H68(n+1)$, and $H2''(n)-H68(n)$ are marked with dashed lines. **B:** Drawings of a DNA dinucleotide in BI and BII conformations, with the hydrogen atoms omitted for clarity.







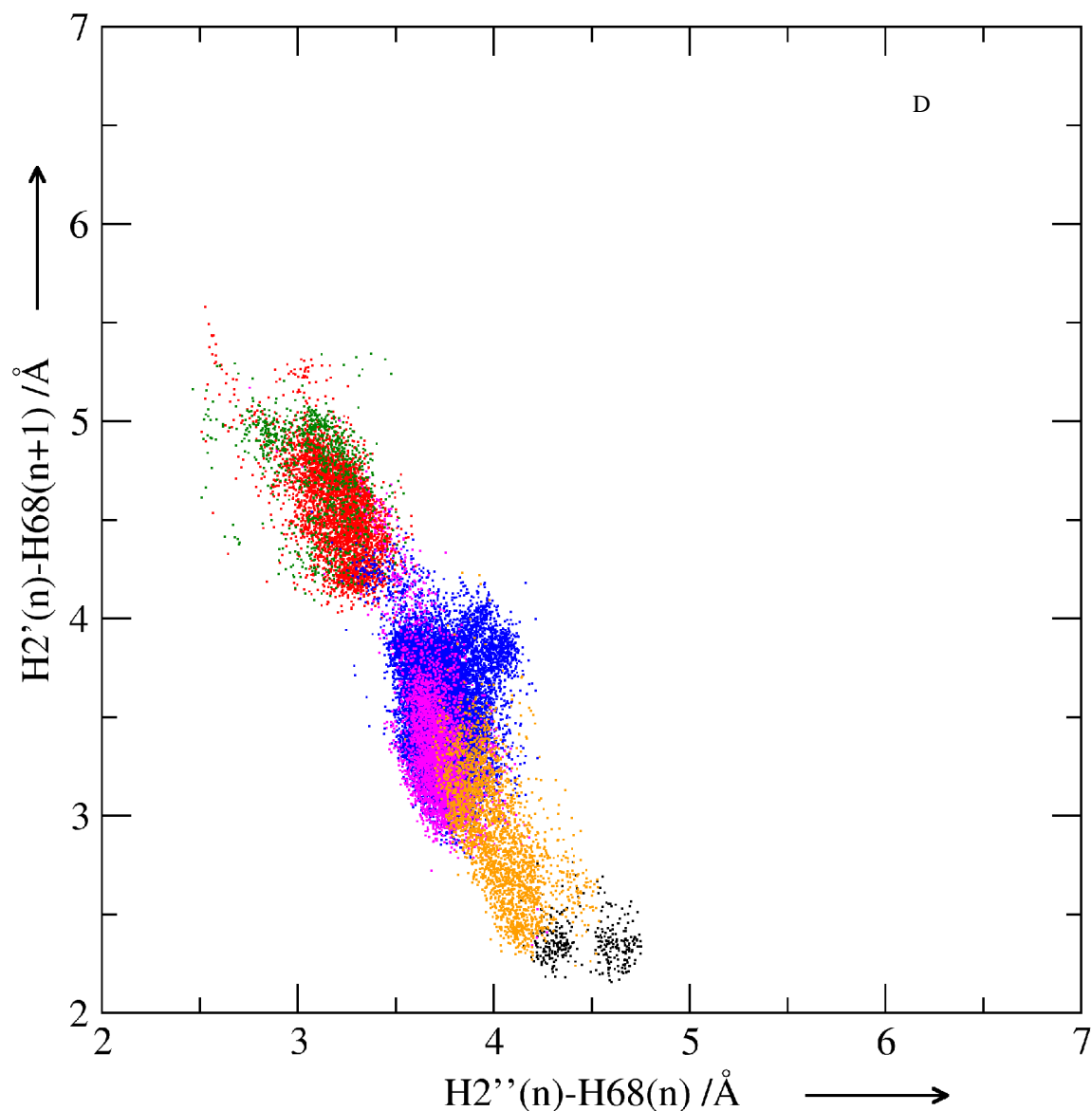


Figure 2A-D. Two-dimensional plots of interproton distances occurring in the ensemble of 26978 dinucleotide steps from four MD simulations. Each point represents one dinucleotide step and the color code indicates its conformation. Blue: **BI**, 5'-sugar S, 3'-phosphate BI. Orange: **BI**, 5'-sugar E, 3'-phosphate BI. Black: **BI**, 5'-sugar N, 3'-phosphate BI. Magenta: **BI**, 3'-phosphate BII. Red: **BII**, face-to-face phosphate BI. Green: **BII**, face-to-face phosphate BII. The sugar pucker assignment was based on the phase angle P : N, $P < 60^\circ$; E, $60^\circ \leq P < 115^\circ$; S, $115^\circ \leq P < 220^\circ$.

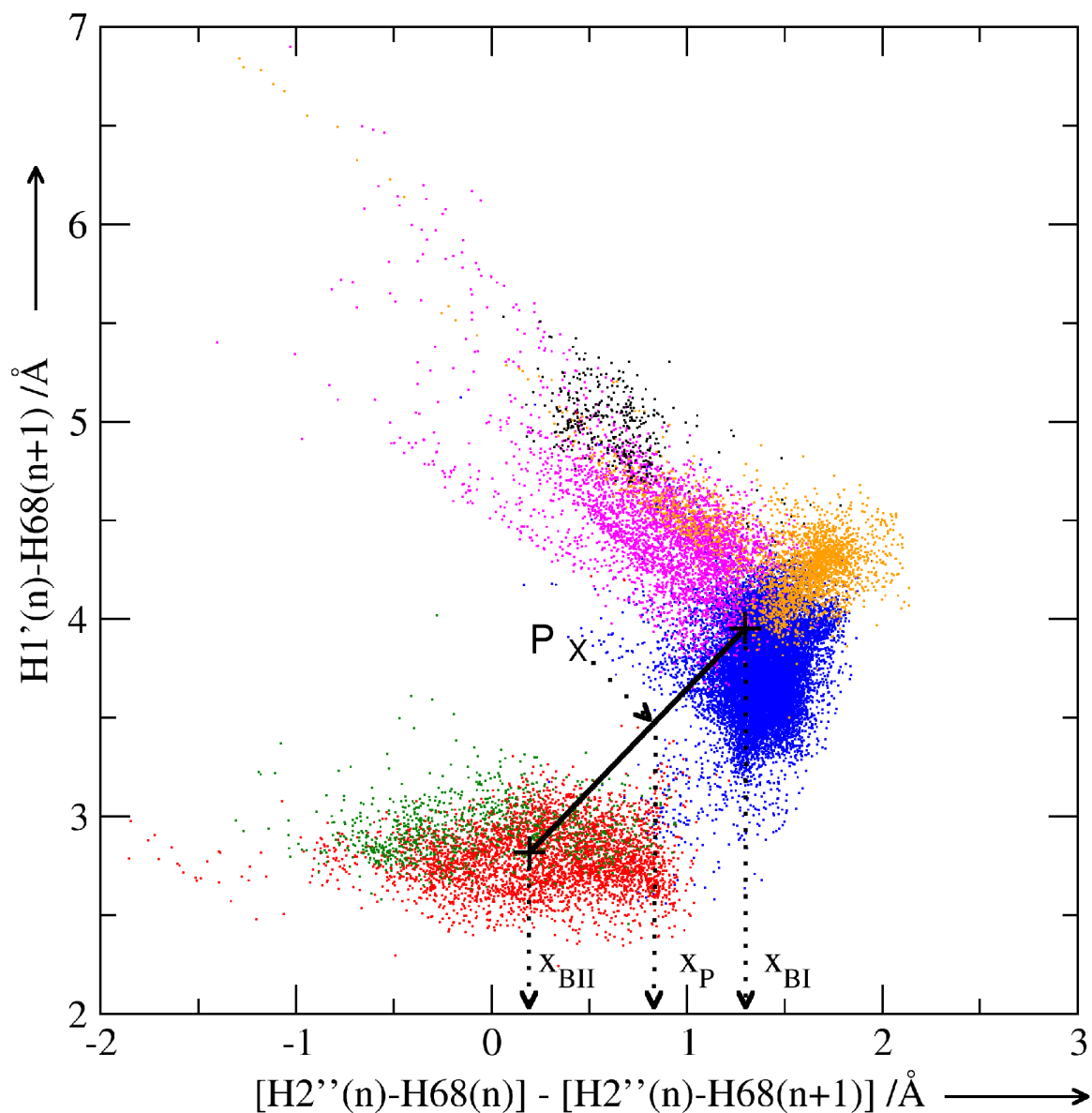


Figure 3. Two-dimensional plot of the difference between distances $H2''(n)-H68(n)$ and $H2''(n)-H68(n+1)$ against the distance $H1'(n)-H68(n+1)$. The ensemble of dinucleotide steps analyzed and the color code are the same as in Figure 2. "+" indicate the centers of gravity of the BI and BII domains. P represents an example of a point that would be determined from the NOE-based measurement of time-averaged distances for a given dinucleotide step. Its projection on the straight line connecting the two centers of gravity yields x_P , which is used in Eq. (1) to calculate % BII.

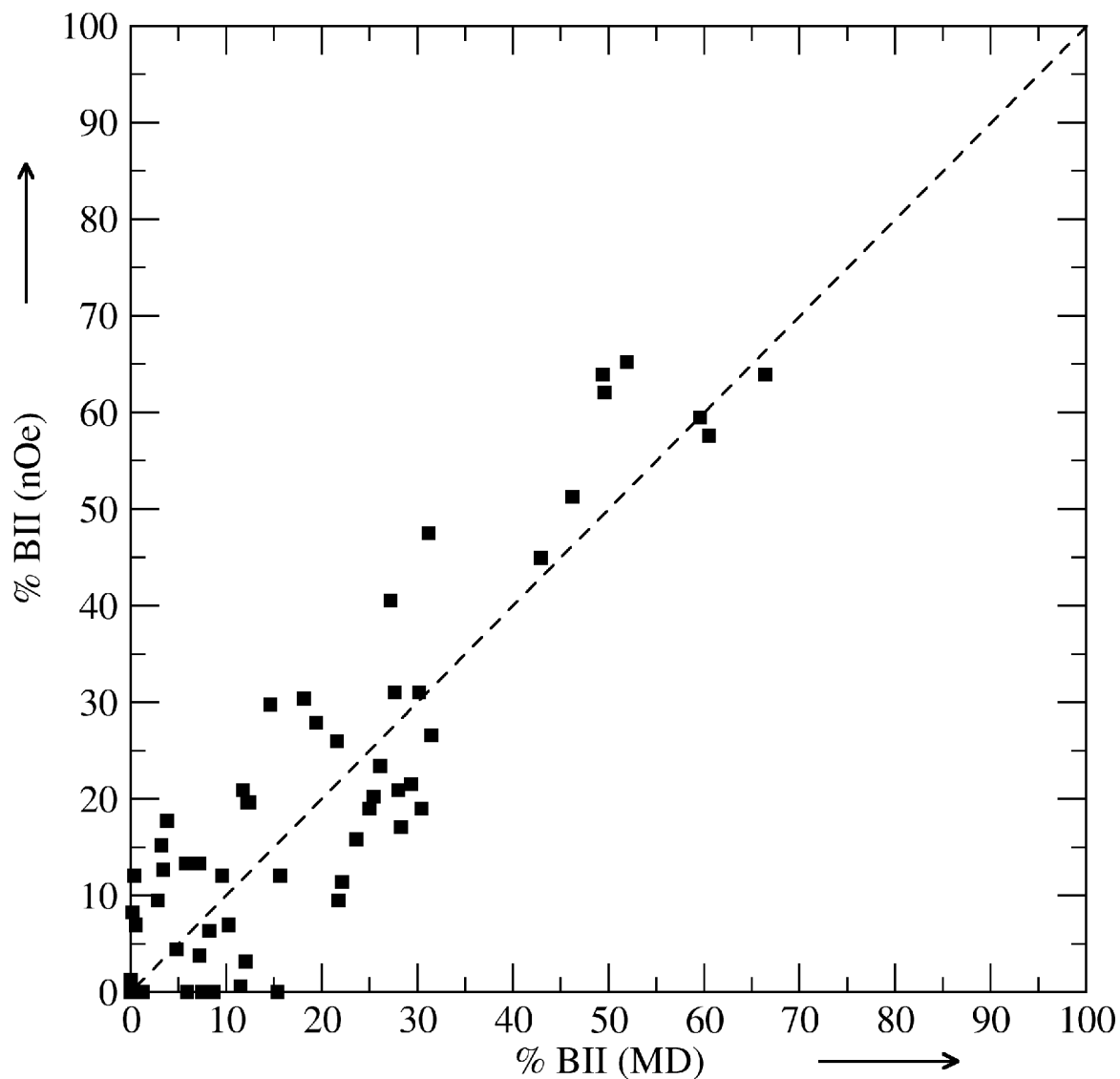


Figure 4. Correlation of the actual BII fractions %BII(MD) occurring in the 64 dinucleotide steps with the "NOE-based" BII fractions %BII(NOE), calculated on the basis of the r^{-6} -averaged distances $H2''(n)$ - $H68(n)$, $H1'(n)$ - $H68(n+1)$, and $H2''(n)$ - $H68(n+1)$, using Figure 3 and Eq. (1) as described in the text. The dashed line corresponds to the equality %BII(NOE) = %BII(MD).

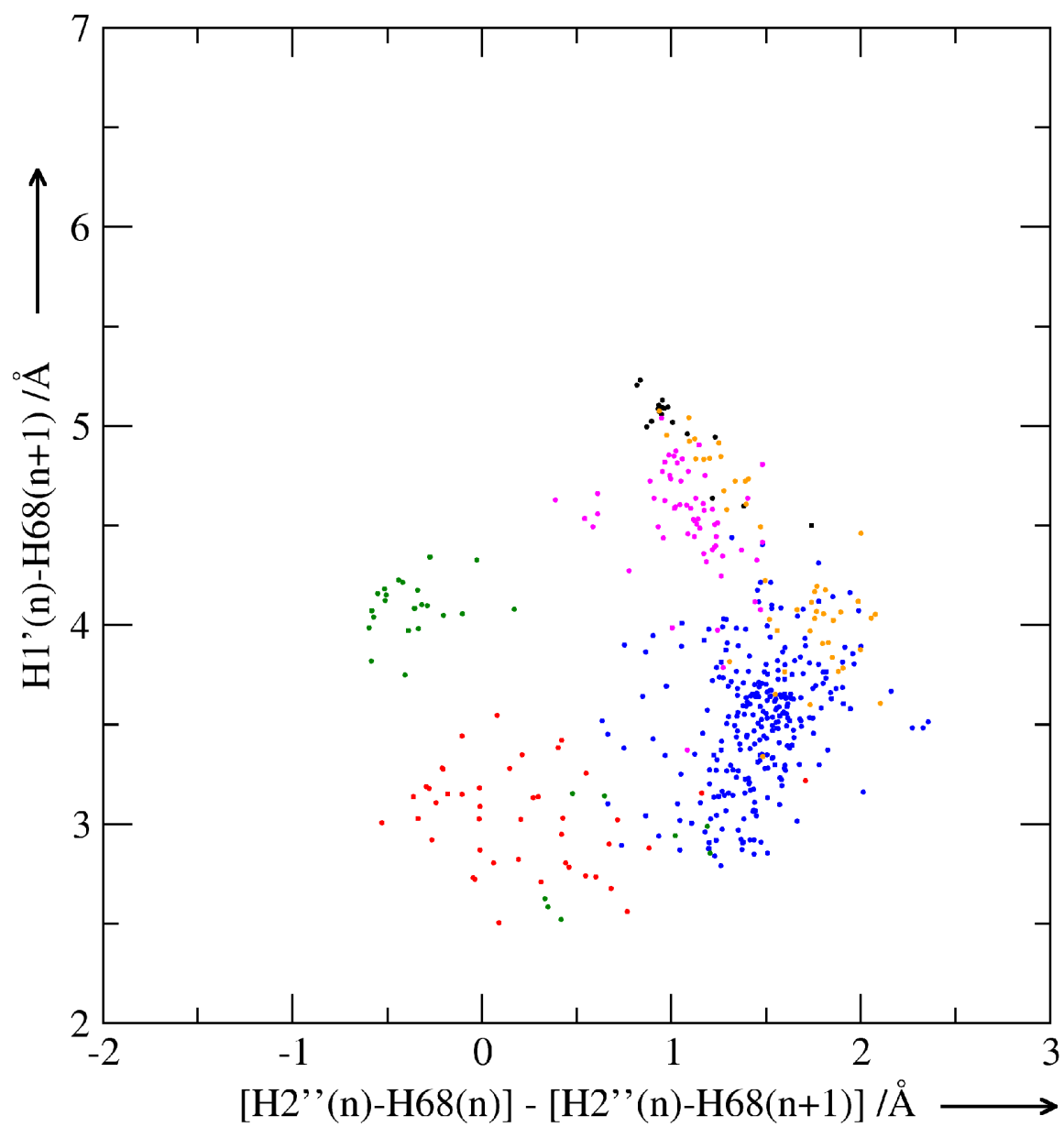


Figure 5. Two-dimensional plot analogous to that in Figure 3, for the ensemble of non-terminal dinucleotide steps of the 39 crystal structures selected from the NDB. See Figure 2 for the color code.

Discrimination between BI and BII conformational substates of B-DNA based on sugar-base interproton distances

Stéphane Teletchéa, Brigitte Hartmann, and Jiří Kozelka

SUPPLEMENTARY MATERIAL

Table S1. Sequences and references of the 39 crystal structures selected from the NDB for analysis of dinucleotide steps.

NDB code	Sequence	Resolution (Å)	Ref.
bd0001	ACCGACGTCGGT	1,60	PROC.NAT.ACAD.SCI.USA V. 95 15194 1998
bd0005	CGCGAATTCGCG	1,75	BIOCHEMISTRY V. 37 16877 1998
bd0006_alpha *	CCAATTGG	1,15	ACTA CRYSTALLOGR., SECT.D V. 55 1495 1999
bd0006_beta *	CCAATTGG	1,15	ACTA CRYSTALLOGR., SECT.D V. 55 1495 1999
bd0007	CGCGAATTCGCG	1,10	J.AM.CHEM.SOC. V. 121 470 1999
bd0012	CGCGAATTCGCG	1,20	J.AM.CHEM.SOC. V. 121 3590 1999
bd0013	CGCGAATTCGCG	1,50	J.AM.CHEM.SOC. V. 121 3590 1999
bd0014	GCGAATTCGC	1,45	J.BIOL.CHEM. V. 274 24749 1999
bd0018	GCGAATTCGC	1,30	J.MOL.BIOL. V. 291 83 1999
bd0019	GCGAATTCGC	1,70	J.MOL.BIOL. V. 291 81 1999
bd0023	CCAGTACTGG	0,74	J.MOL.BIOL. V. 296 787 2000
bd0029	CGCGAATTCGCG	1,82	J.AM.CHEM.SOC. V. 122 1546 2000
bd0030_alpha *	CGCGAATTCGCG	0,95	BIOPOLYMERS V. 48 234 1998
bd0030_beta *	CGCGAATTCGCG	0,95	BIOPOLYMERS V. 48 234 1998
bd0033	CCAACGTTGG	1,00	J.MOL.BIOL. V. 301 915 2000
bd0034	CCAACGTTGG	1,00	J.MOL.BIOL. V. 301 915 2000
bd0035	CCAGCGCTGG	1,00	J.MOL.BIOL. V. 301 915 2000
bd0036	CCAGCGCTGG	1,00	J.MOL.BIOL. V. 301 915 2000
bd0037	CGAATTCG	0,89	J.BIOL.CHEM. V. 275 23034 2000
bd0038	CGCGAATTCGCG	1,43	BIOCHEMISTRY V. 39 3525 2000
bd0041	CGCGAATTCGCG	1,20	J.AM.CHEM.SOC. V. 122 11048 2000
bdf068	CTCGAG	1,90	BIOPHYS.J. V. 70 2857 1996
bdj017	CCAGGCCTGG	1,60	J.Mol.Biol. V. 210 369 1989
bdj019	CCAACGTTGG	1,40	J.MOL.BIOL. V. 217 177 1991
bdj025	CGATCGATCG	1,50	J.BIOL.CHEM. V. 266 8861 1991
bdj031	CGATTAATCG	1,50	J.MOL.BIOL. V. 225 379 1992
bdj036	CGATATATCG	1,70	BIOCHEMISTRY V. 31 8009 1992
bdj037	CGATATATCG	2,00	BIOCHEMISTRY V. 31 8009 1992
bdj051	CATGGCCATG	2,00	PROC.NAT.ACAD.SCI.USA V. 90 2930 1993
bdj052	CCAAGCTTGG	1,90	BIOCHEMISTRY V. 32 8923 1993
bdj060	CTCTCGAGAG	1,70	BIOCHEMISTRY V. 34 1022 1995
bdj061	CCACTAGTGG	1,95	NATURE V. 368 469 1994
bdj081_AB *	CAAAGAAAAG	1,85	J.MOL.BIOL. V. 269 811 1997
bdj081_CD *	CAAAGAAAAG	1,85	J.MOL.BIOL. V. 269 811 1997
bdj081_EF *	CAAAGAAAAG	1,85	J.MOL.BIOL. V. 269 811 1997
bdl001	CGCGAATTCGCG	1,90	PROC.NAT.ACAD.SCI.USA V. 78 2179 1981
bdl005	CGCGAATTCGCG	1,90	ACTA CRYSTALLOGR.,SECT.B V. 41 255 1985
bdl020	CGCGAATTCGCG	1,90	J.BIOMOL.STRUCT.DYN. V. 5 581 1987
bdl084	CGCGAATTCGCG	1,40	BIOCHEMISTRY V. 37 8341 1998

* crystals containing multiple conformations

Deuxième partie

RMN de l'adduit platiné G^*G^*G

L'activité antitumorale du cisplatine ($\text{cis-[PtCl}_2(\text{NH}_3)_2]$) a fait l'objet de nombreuses études structurales sur les adduits qu'il forme avec l'ADN, sa cible cellulaire principale. L'adduit majoritaire, nommé adduit GG intrabrin, est formé par la coordination du cisplatine sur deux guanines adjacentes du même brin. Sa structure a été étudiée par de nombreuses équipes [49, 56, 71, 87, 129, 210]. Toutes les séquences étudiées à ce jour sont de la forme $\text{pyG}^*\text{G}^*\text{N}$, où py = pyrimidine et $\text{N} = \text{T, C}$, ou A (passé en revue dans [4, 58]).

Un oligonucléotide double-brin contenant la séquence GGG n'a pas encore fait l'objet d'une étude structurale, alors que les séquences contenant G_n ($n \geq 3$) sont des points chauds pour la platination [57].

Une des raisons qui a probablement dissuadé les membres des équipes travaillant sur le cisplatine est la difficulté de purification des adduits formés. En effet, la séquence GGG réagit avec le cisplatine pour former des adduits 1,2-GG et 2,3-GG et dans une moindre mesure 1,3-GG [57], les deux premiers étant très difficiles à séparer l'un de l'autre. Nous avons développé à cet effet une méthode de purification en HPLC inédite qui nous a permis de séparer et purifier les adduits 1,2-GG des adduits 2,3-GG. Cette méthode a été mise au point en s'appuyant sur l'expérience acquise au laboratoire dans les études cinétiques de platination et chélation ADN-cisplatine [137].

Nous avons choisi la séquence $\text{d}[(\text{GCCGGGTTCGC}) \cdot (\text{GCGACCCGGC})]$ **(1)** (noté GGG, ou $\text{G}^*\text{G}^*\text{G}/\text{GG}^*\text{G}^*$ si la séquence est platinée, * représente la base cible de la coordination) où la paire A_6T_{15} de la séquence $\text{d}[(\text{GCCG}^*\text{G}^*\text{ATTCGC}) \cdot (\text{GCGATCCGGC})]$ **(2)** (noté $\text{G}^*\text{G}^*\text{A}$) étudiée précédemment [59, 87] a été échangée pour la paire G_6C_{15} . Cela a permis la comparaison entre les adduits G_4^*G_5^* de **(1)** et **(2)**, en particulier l'étude de la transition $\text{A} \rightarrow \text{G}$ en position 3' de l'adduit. Une telle transition a été prédite par simulation pour stabiliser la conformation BII sur le pas G_6pT_7 .

La platination a été réalisée sur le brin GCCGGGTTCGC en utilisant le cisplatine dont les ammines sont marquées ($\text{cis-[PtCl}_2(^{15}\text{NH}_3)_2]$). Les deux adduits majeurs ont été séparés et purifiés par HPLC, l'adduit principal $\text{G}_4^*\text{G}_5^*\text{G}_6$ a été hybridé avec son brin complémentaire. Le duplex a été étudié par RMN et simulation de dynamique moléculaire.

Ce chapitre présente l'analyse des données spectroscopiques RMN, la confrontation a posteriori avec les simulations de dynamique moléculaire et la première observation jamais décrite du réarrangement de l'adduit $\text{G}_4^*\text{G}_5^*\text{G}_6$ vers l'adduit $\text{G}_4\text{G}_5^*\text{G}_6^*$.

Chapitre 1

Préparation de l'adduit platiné

La séquence GGG réagit avec le cisplatine pour donner lieu majoritairement aux pontages intra-brin 1,2 et 2,3. Dans une étude de la reconnaissance des duplex d'ADN platinés par la protéine LEF-1, à laquelle j'ai initialement participé, le Dr. Drumm a fait réagir le 20-mer TCTCCTTCCGGGTTCTTCTC avec le cisplatine. Nous avons mis au point la séparation par HPLC en phase inverse du mélange réactionnel et isolé deux fractions majeures que nous avons pu identifier comme les adduits 1,2 G*G*G et 2,3 GG*G* en utilisant le séquençage de type Maxam-Gilbert. Le même protocole a été appliqué à la séparation du mélange réactionnel entre le décanucléotide GCCGGGTCGC et le cisplatine.

Pour la préparation du décanucléotide en concentration suffisante pour la RMN, le simple brin 5'-GCCGGGTCGC-3' (GGG) à la concentration initiale de $2.10^{-4}\text{mol.l}^{-1}$ a été platiné en le faisant réagir avec le cisplatine dichloro cis-[PtCl₂(¹⁵NH₃)₂] à l'abri de la lumière, à une température de 37°C et un pH ajusté à 4,5 par HClO₄. Le volume total du bain réactionnel a été fixé à 25 ml afin d'obtenir une quantité de 7,5 micromoles du mono-brin platiné.

La réaction a été initiée avec 1 équivalent de cisplatine, puis progressivement, par ajout de 1+0,5+0,5 équivalent, amenée à 3 équivalents pour une durée totale de réaction de 51 heures. La réaction a été suivie par HPLC (figure 1.1) et interrompue quand les produits éluants avant l'oligonucléotide de départ (correspondant à des mono-adduits puis probablement à des adduits multiplatinés) ont commencé à progresser plus vite que les deux chélates intrabrin majeurs. La figure 1.1 présente le déroulement de la platination et l'apparition des adduits en fonction de l'évolution de la réaction. Au début de la platination, tous les mono adduits apparaissent (G₁*, G₄*, G₅*, G₆* et G₉*) mais les mono-adduits G₄*, G₅*, G₆* peuvent se réarranger en chélates tandis que les mono adduits G₁* et G₉* ne peuvent se réarranger et s'accumulent. Seuls les deux chélates G₄*G₅*G₆ (G*G*G) et G₄G₅*G₆* (GG*G*) ont été séparés et caractérisés par le séquençage de type Maxam-Gilbert.

Les fractions contenant les adduits G^*G^*G et GG^*G^* ont été collectées, dialysées, puis lyophilisées. Après estimation par densité optique (coefficient d'extinction molaire $\varepsilon = 88300 \text{ l.mol}^{-1}.\text{cm}^{-1}$ [25]) suivant la loi de Beer-Lambert de la concentration résiduelle des fractions, la fraction contenant le chélate $G_4^*G_5^*G_6$ a été hybridée avec le brin complémentaire ($\varepsilon = 88900 \text{ l.mol}^{-1}.\text{cm}^{-1}$) pour former le duplex final.

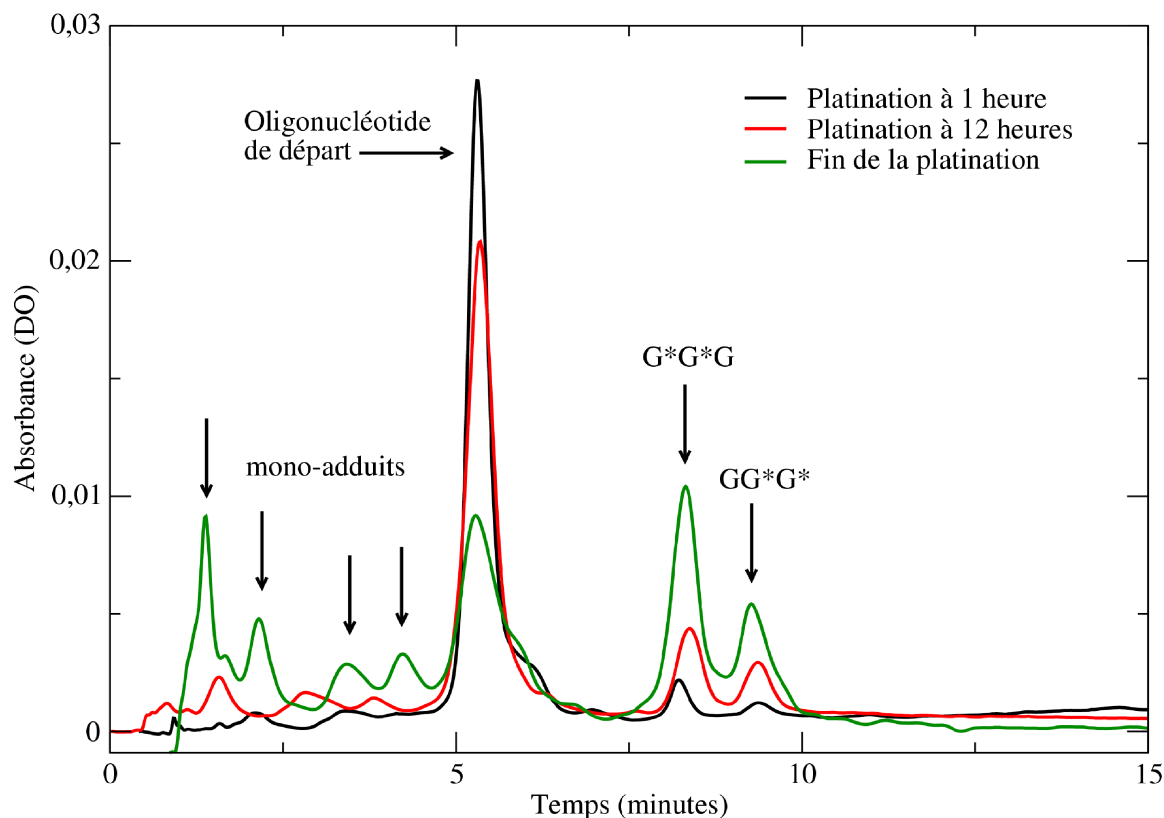


FIG. 1.1 – Evolution de la réaction du décanucléotide GCCGGGTTCGC avec le cisplatine ($\text{cis-}[\text{PtCl}_2(^{15}\text{NH}_3)_2]$). 1 équivalent du cisplatine a été ajouté au temps $t=0$, puis addition de 1 équivalent à 20h puis 0,5 équivalent à 36h et 43h. Le gradient HPLC utilisé commence à 2% d'acétonitrile puis augmente de manière exponentielle à 20 % jusqu'à 25 minutes (programme 3 du logiciel Waters Millenium32 ver. 3.05.01). Les chélates sont séparés en HPLC au bout de 10 minutes, soit un pourcentage d'acétonitrile de 8%. Le passage de 8 à 20 % permet de rincer la colonne pour effectuer une nouvelle analyse. Le solvant organique est constitué d'acétonitrile (mais est dilué dans 50% d'eau pour faciliter l'analyse), le solvant aqueux contient l'acétate d'ammonium 0,5 M ajusté à pH 4,5.

Chapitre 2

RMN de l'adduit

Le spectrographe de Résonance Magnétique Nucléaire de type Bruker BX-600 a été utilisé. En raison de la quantité faible concentration de la fraction G*G*G (250 $\mu\text{mol.l}^{-1}$), un tube de type Shigemi (Sigma-Aldrich) a servi pour l'analyse, permettant d'atteindre une concentration apparente de 0,5 mM dans le tube. La fraction GG*G* n'a quant à elle pas pu être utilisée, la purification de l'échantillon n'ayant pas permis d'obtenir une concentration suffisante pour l'expérience RMN.

2.1 Analyse de l'échantillon

L'utilisation combinée des expériences NOESY, COSY et TOCSY en D_2O [169, 204, 205] a permis l'identification de tous les protons non échangeables de l'oligonucléotide non platiné d[(GCCGGGTCGC)·(GCGACCCGGC)] (GGG) à l'exception de H5'/H5". Pour l'oligonucléotide platiné [(GCCG*G*GTCGC)·(GCGACCCGGC)] (G*G*G) les protons (x)H5'/H5", G1H4', G11H4' et G13H4' et G19H4' n'ont pas pu être assignés. L'ensemble des attributions est présenté dans les tableaux 2.2 et 2.1.

Les protons imino échangeables ont été attribués à partir des spectres enregistrés à 281 K de l'échantillon repris dans 90/10 $\text{H}_2\text{O}/\text{D}_2\text{O}$ en suivant les connectivités des amines des cytosines, du proton H2 de l'adénine avec respectivement leurs guanines et sa thymine.

Pour G*G*G les connectivités séquentielles pour H1'/H2'/H2''(n)-H6/8(n+1) ont pu être observées pour les deux brins, à l'exception du site de platination où les signaux d'interaction pour C3H1'-G4*H8 et G*4H1'-G*5H8 sont manquants (voir tableau 2.1, figure 2.1 et figure 2.4). Ces observations, similaires à celles faites pour dsG*G*A, sont caractéristiques des adduits ADN(pyGG)-Pt. L'absence du signal s'explique par le désempilement entre les deux bases platinées et entre la base platinée 5' et la pyrimidine qui la précède [87].

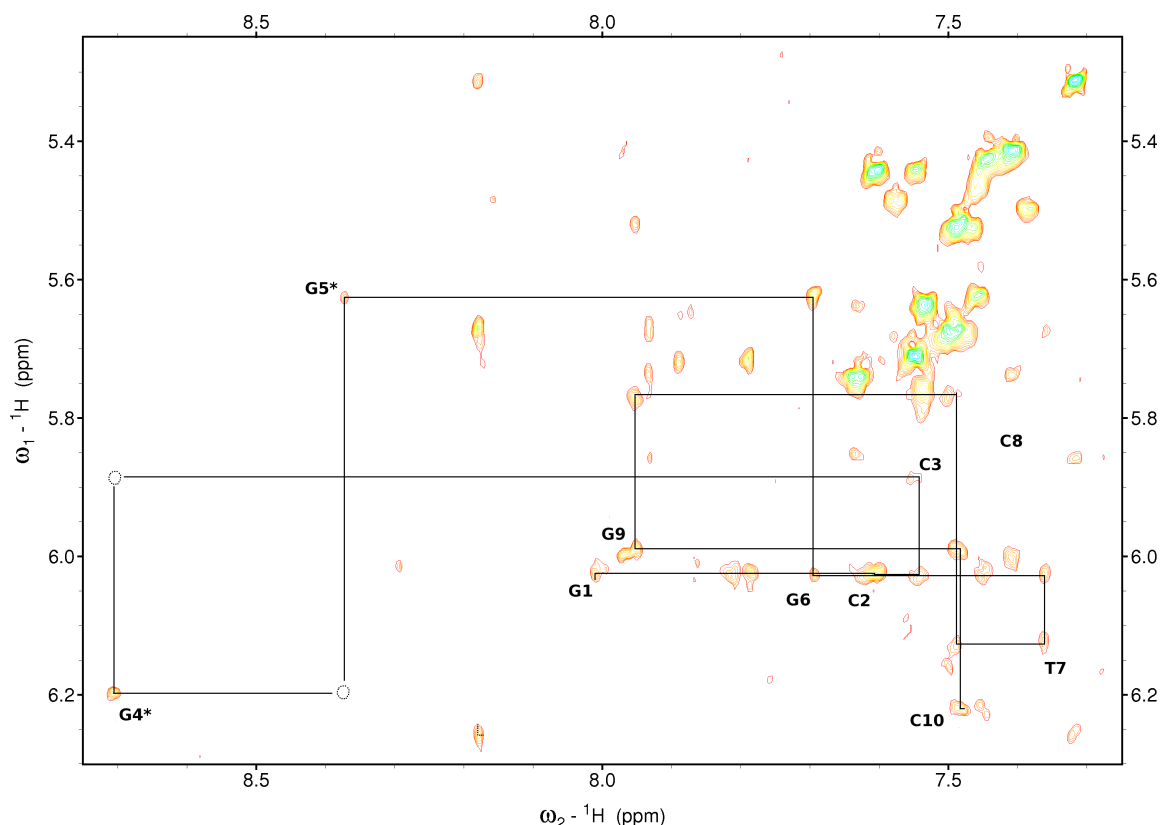


FIG. 2.1 – Vue étendue du spectre NOESY de la région G*G*G, d[(GCCG*G*GTCGC) · (GCGACCCGGC)]. La vue présente le chemin d'attribution du brin platiné dans la région aromatique-anomérique. Les cercles en pointillés indiquent les pics manquants. Les pics qui ne font pas partie du chemin d'attribution appartiennent au brin complémentaire et au duplex GG*G*. La concentration de l'oligonucléotide est de 0,50 mM dans 1 M NaClO₄, 30 mM tampon phosphate, pH 6,0 en D₂O à 305 K.

2.2 Déplacements chimiques observés

Le désempilement et le basculement entre les bases platinées dans l'orientation tête-tête se manifeste par un signal NOE H8-H8 entre ces bases platinées d'intensité moyenne. Un signal NOE très faible entre G*4H1-G*5H8 est aussi l'indicateur du désempilement. Toutes les connectivités concernant la région H2'/H2''-H6/H8, en dépit du désempilement, sont en conformité avec celles observées pour G*G*A.

Les caractéristiques communes entre le spectre 2D NOESY de G*G*G et les complexes pyG*G* sont :

- un signal faible pour H1'(n)-H8(n+1) et H3'(n)-H8(n+1) entre la cytidine complémentaire de la guanine 5' platinée (C17) et la purine suivante (G18) ;
- un signal H3'(n)-H8(n+1) manquant entre C3 et G*4.

Le signal C16H1'-C17H6 entre les deux cytidines complémentaires des guanines platinées est observé avec une intensité moyenne alors que dans d'autres études le pic a

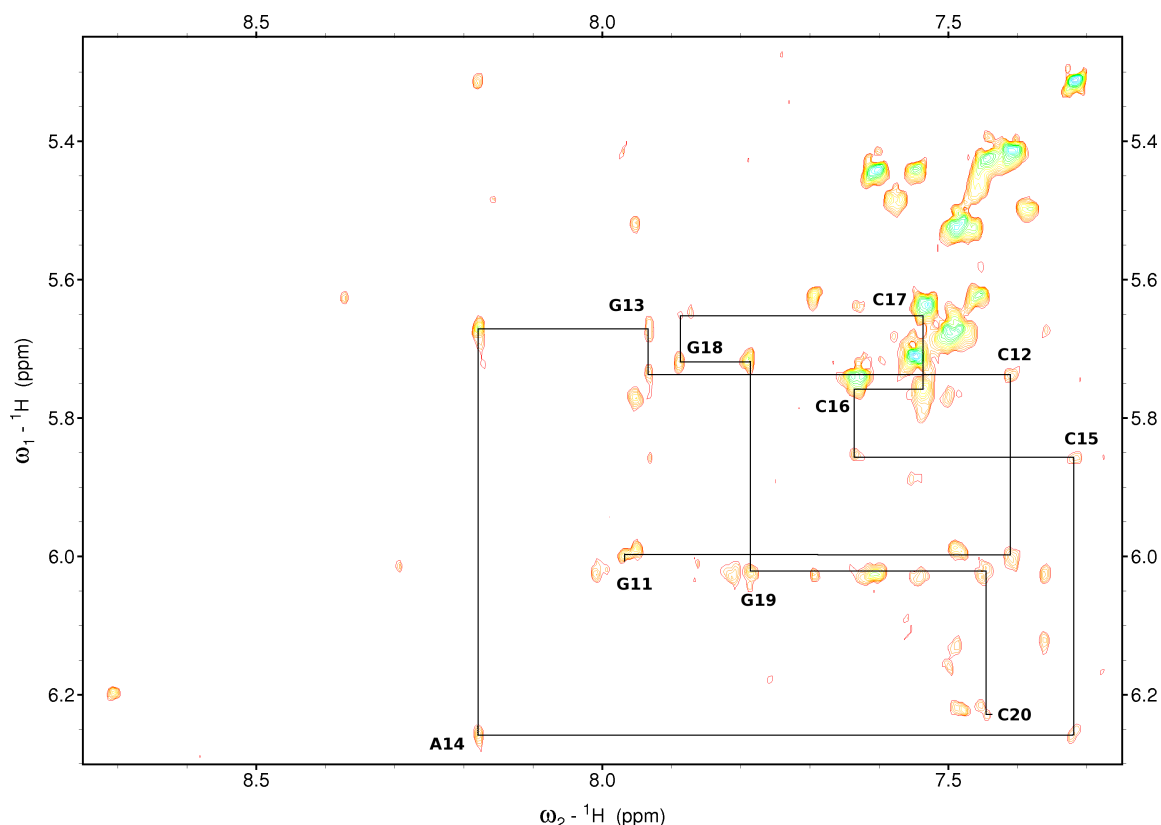


FIG. 2.2 – Vue étendue du spectre NOESY de la région G*G*G, d[(GCCG*G*GTCGC) · (GCGACCCGGC)]. La vue présente le chemin d’attribution du brin complémentaire non platiné dans la région aromatique-anomérique. Les pics qui ne font pas partie du chemin d’attribution appartiennent au brin platiné et au duplex GG*G*. La concentration de l’oligonucléotide est de 0,50 mM dans 1 M NaClO₄, 30 mM tampon phosphate, pH 6,0 en D₂O à 305 K.

parfois une intensité moyenne [71, 210], parfois une intensité faible ou est manquant [87]. Cela peut venir de la variabilité que présente la phase des sucres sur ces cytidines (présentée dans [4], p. 272). La distance H1’(n)-H8(n+1) est sensible à la phase du sucre du nucléotide n, elle est plus longue quand le sucre est en phase N [185]. Il semble donc que C16 ait un caractère plus N dans le cas de G*G*A [87] que pour G*G*G (voir ci-dessous).

A basse température (281 K) les couplages imino-imino T7H3-G6H1 et T7H3-G13H1 sont présents, tandis que le couplage G4H1-G5H1 est manquant et le couplage G5H1-G6H1 est très faible.

Les tableaux 2.1 et 2.2 indiquent les déplacements chimiques de l’oligonucléotide platiné G*G*G et de l’oligonucléotide non platiné. Le duplex platiné présente toutes les caractéristiques des complexes pyG*G* référencées dans [4]. En particulier, les deux protons H8 des guanines platinées sont déblindés. Leurs déplacements chimiques sont caractéristiques des complexes pyG*G* : 8,7 ppm pour la guanine platinée 5’ et 8,0-8,4

ppm pour la guanine platinée 3'. Le proton H2' de la guanine platinée en 5' et celui de la cytidine complémentaire sont blindés comme le sont ceux de la majorité des adduits pyG*G* [59] et résonnent respectivement à 1,52 et 1,98 ppm (tableau 2.1, à corrélérer avec la figure 13 de [59]). C'est un indicateur que la structure de G*G*G possède les caractéristiques communes des autres structures de pyG*G* étudiées jusqu'à présent [49, 56, 71, 87, 129, 210].

	H8/H6	H2/H5/CH3	H1'	H2'	H2''	H3'	H4'	H1/H3	H41	H42
G1	8,01	-	6,02	2,74	2,79	4,85	-	-	-	-
C2	7,61	5,45	6,03	2,30	2,50	4,84	4,26	-	8,37	6,48
C3	7,55	5,71	5,89	1,52	2,46	4,71	4,10	-	8,85	7,11
G4*	8,71	-	6,20	2,33	2,75	5,17	4,25	13,16	-	-
G5*	8,37	-	5,63	2,41	2,66	4,83	4,25	-	-	-
G6	7,70	-	6,03	2,54	2,78	4,88	4,38	12,93	-	-
T7	7,36	1,30	6,13	2,21	2,56	4,92	4,28	13,79	-	-
C8	7,49	5,67	5,77	2,07	2,43	4,87	4,15	-	8,55	6,87
G9	7,95	-	5,99	2,66	2,76	5,01	4,40	13,10	-	-
C10	7,49	5,52	6,23	2,19	2,21	4,53	4,08	-	-	-
	H8/H6	H2/H5/CH3	H1'	H2'	H2''	H3'	H4'	H1/H3	H41	H42
G11	7,97	-	6,00	2,60	2,80	4,87	-	-	-	-
C12	7,41	5,42	5,74	2,06	2,43	4,89	4,20	-	8,52	-
G13	7,93	-	5,67	2,77	2,84	5,06	-	12,92	-	-
A14	8,18	7,93	6,26	2,77	2,89	5,03	4,49	-	-	-
C15	7,32	5,32	5,86	2,08	2,48	4,75	4,16	-	8,31	6,74
C16	7,63	5,74	5,76	2,06	2,48	4,82	4,15	-	8,61	7,04
C17	7,53	5,64	5,65	1,98	2,31	4,82	4,07	-	8,66	6,83
G18	7,89	-	5,72	2,72	2,79	5,00	4,41	12,98	-	-
G19	7,79	-	6,03	2,59	2,78	5,00	-	-	-	-
C20	7,45	5,43	6,22	2,21	2,22	4,53	4,08	-	-	-

TAB. 2.1 – Déplacements chimiques pour ^1H du duplex platiné G*G*G. Tous les spectres ont été enregistrés à 305 K à l'exception des spectres des protons échangeables pour G*G*G enregistrés à 281 K. Les déplacements ont pour référence le TSP (Trisilyl Methyl Phosphate).

La comparaison entre les déplacements chimiques des protons non échangeables de G*G*G et G*G*A (figure 2.3) suggère qu'une grande partie des deux duplexes est surperposable. Des différences significatives sont présentes entre G6/A6 et sur C16. La transition de A vers G explique les différences observées dans les déplacements chimiques des protons du sucre [190]. Le changement de T17 pour C17 induit un effet de courant de cycle différent [73] qui explique les changements de déplacements chimiques. La ressemblance structurale entre G*G*G et G*G*A est de plus renforcée par les signaux observés dans le spectre RMN de ^{31}P .

	H8/H6	H2/H5/CH3	H1'	H2'	H2''	H3'	H4'	H1/H3	H41	H42
G1	7,98	-	6,01	2,65	2,78	4,85	4,26	-	-	-
C2	7,53	5,42	6,09	2,18	2,49	4,88	4,27	-	8,34	6,45
C3	7,42	5,60	5,62	1,99	2,35	4,84	4,11	-	8,62	6,75
G4	7,83	-	5,64	2,66	2,76	4,99	4,34	13,07	-	-
G5	7,67	-	5,81	2,64	2,74	4,96	4,39	12,95	-	-
G6	7,53	-	5,92	2,47	2,73	4,83	4,37	12,82	-	-
T7	7,28	1,26	6,05	2,12	2,52	4,86	4,24	13,65	-	-
C8	7,47	5,64	5,73	2,04	2,41	4,85	4,13	-	8,53	6,83
G9	7,93	-	5,97	2,63	2,74	4,99	4,37	13,08	-	-
C10	7,46	5,48	6,20	2,19	2,18	4,50	4,06	-	-	-
	H8/H6	H2/H5/CH3	H1'	H2'	H2''	H3'	H4'	H1/H3	H41	H42
G11	7,96	-	5,99	2,60	2,78	4,85	4,24	-	-	-
C12	7,39	5,40	5,71	2,02	2,39	4,87	4,17	-	8,49	6,48
G13	7,91	-	5,61	2,72	2,81	5,03	4,36	12,87	-	-
A14	8,16	7,87	6,25	2,70	2,91	5,03	4,48	-	-	-
C15	7,23	5,24	5,79	2,03	2,42	4,74	4,18	-	8,02	6,48
C16	7,44	5,44	5,90	2,08	2,42	4,80	4,14	-	8,34	6,55
C17	7,40	5,57	5,57	1,98	2,34	4,82	4,09	-	8,59	6,74
G18	7,85	-	5,65	2,68	2,76	4,99	4,34	13,16	-	-
G19	7,76	-	5,99	2,54	2,74	4,97	4,38	13,12	-	-
C20	7,43	5,42	6,18	2,17	2,19	4,50	4,04	-	-	-

TAB. 2.2 – Déplacements chimiques pour ^1H du duplex non-platiné GGG. Tous les spectres ont été enregistrés à 305 K. Les déplacements ont pour référence le TSP (Trisilyl Méthyl Phosphate).

2.3 Différences entre platiné et non platiné

La figure 2.4 présente les modifications observées suite à la platination de l'oligonucléotide. Il apparaît clairement que les changements observés sont localisés au niveau du site de platination (C_3 , G_4^* , G_5^* , A_6) et sur les bases complémentaires des guanines platinées (C_{15} , C_{16}). La déformation engendrée par la coordination du cisplatine est très limitée.

Les différences de déplacements chimiques sont moins importantes que celles résultant de la transition $\text{A} \rightarrow \text{G}$ entre $\text{G}^*\text{G}^*\text{G}$ et $\text{G}^*\text{G}^*\text{A}$ (figure 2.3).

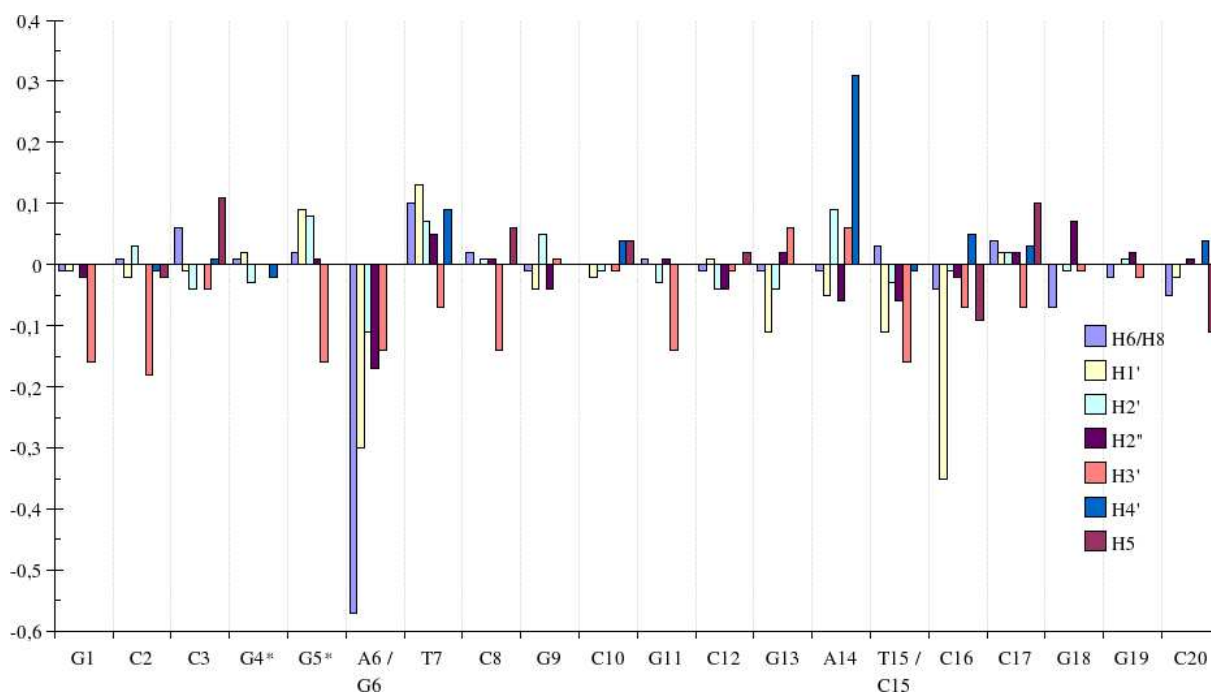


FIG. 2.3 – Différence entre les déplacements chimiques pour les protons non échangeables entre G*G*G et G*G*A. Les spectres ont été enregistrés à 305K (G*G*G) et 298K (G*G*A).

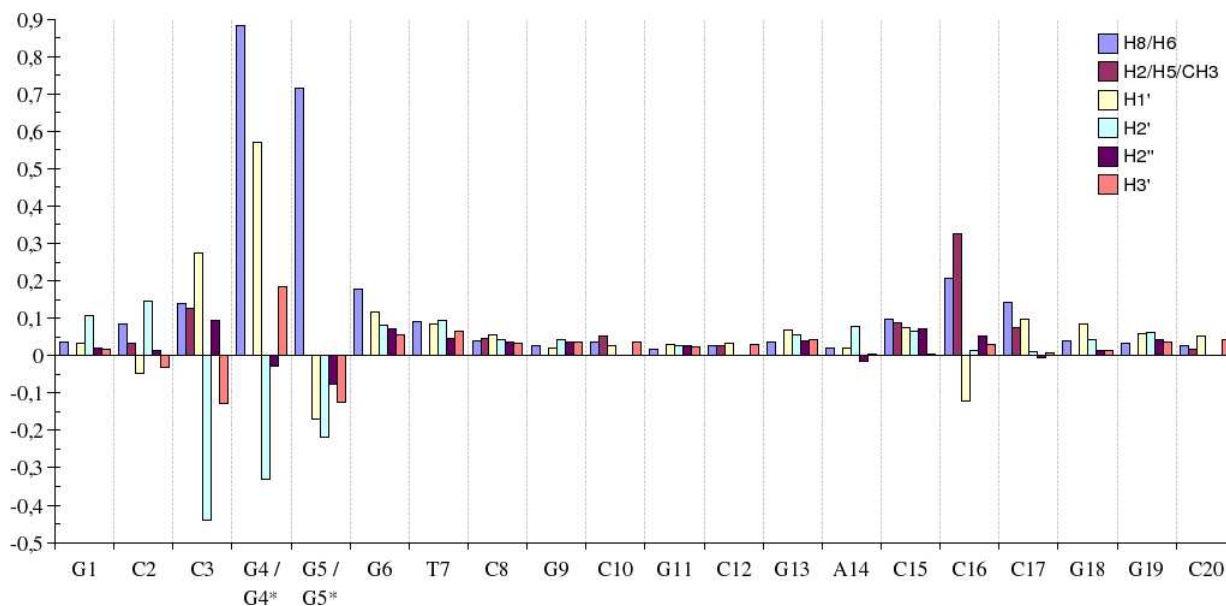


FIG. 2.4 – Différence entre les déplacements chimiques pour les protons non échangeables entre G*G*G et GGG. Les spectres ont été enregistrés à 305K (G*G*G).

2.4 Attribution des sucres

2.4.1 Phase des sucres

La coordination du cisplatine est connue pour altérer la phase des sucres des nucléotides platinés [4]. Dans tous les complexes pyG*G* étudiés, la guanine en 5' présente

une phase quasi exclusivement N. Le changement de phase vers N a été observé partiellement [87] ou complètement [129], l'une des cytidines complémentaires des bases platinées présentant aussi un caractère N renforcé (passé en revue dans [4] p. 272). Pour valider ces observations, nous avons porté une attention toute particulière au trinuécléotide central (CG*G*)-(CCG).

2.4.2 Informations issues des constantes de couplages $J_{1'2'}$ et $J_{1'2''}$

Les spectres DQF-COSY ont été enregistrés en utilisant une sonde cryogénique qui a permis la mesure de presque tous les couplages $J_{1'2'}$ et $J_{1'2''}$, avec une précision de ± 0.6 -1.0 Hz (tableau 2.3). La plupart des nucléotides ont un couplage $J_{1'2'}$ de 10 Hz et un couplage $J_{1'2''}$ de 5 Hz, caractéristiques d'une phase S dans l'ADN B [160]. G4* présente un petit couplage $J_{1'2'}$ de 1,9 Hz et un grand couplage $J_{1'2''}$ de 7,1 Hz, compatibles avec une phase de sucre 100% N [160]. Des valeurs intermédiaires sont observées pour C3 et C15.

Dans le cas de C3, les constantes de couplage sont $J_{1'2'}=4,5$ Hz et $J_{1'2''}=8.0$ Hz, similaires à celles observées pour G*G*A (4,5 Hz et 6,5 Hz [87]). La valeur $J_{1'2''}$ plus importante pour G*G*G est l'indicateur que l'équilibre est plus déplacé vers la conformation N dans le cas de G*G*G que pour G*G*A. La simulation de dynamique moléculaire (voir ci-après) indique une conformation majoritairement N avec un angle moyen pour la phase de 25° et une amplitude de Φ_m de 38° , pour lesquels les valeurs de $J_{1'2'}$ et $J_{1'2''}$ sont respectivement de 2,4 Hz et 8,6 Hz, selon Rinkel et Altona [160]. En supposant que les valeurs observées sont le résultat d'un équilibre N/S avec une faible proportion de S ($P=156^\circ$, $\Phi_m=35^\circ$, $J_{1'2'}=10,2$ Hz, $J_{1'2''}=5,7$ Hz [160]), le calcul indique que la proportion 75 % N, 25 % S donne pour les constantes de couplage $J_{1'2'}=4,3$ Hz et $J_{1'2''}=7,9$ Hz. Ces valeurs sont en accord avec les observations expérimentales. Pour G*G*A, la proportion de 60-70 % N pour C3 avait été estimée, pour G*G*G la proportion estimée est donc légèrement plus importante.

Pour C15, la simulation de dynamique moléculaire (voir plus bas) indique des transitions entre S ($P=125^\circ$, $\Phi_m=40^\circ$), N ($P=25^\circ$, $\Phi_m=35^\circ$), et E ($P=75^\circ$, $\Phi_m=38^\circ$). Les constantes de couplages correspondant à ces trois conformations sont $J_{1'2'}=10,5$ Hz, $J_{1'2''}=5,2$ Hz pour S, $J_{1'2'}=7,1$ Hz, $J_{1'2''}=8,2$ Hz pour E et $J_{1'2'}=2,3$ Hz, $J_{1'2''}=8,4$ Hz pour N [160]. En supposant que les proportions des conformations sont 73 % de N, 14% de E et de 13 % pour S, les constantes de couplages deviennent $J_{1'2'}=4,0$ Hz et $J_{1'2''}=8,0$ Hz, en accord avec les observations expérimentales.

Base	$J_{1'2'}$	$J_{1'2''}$	$\Sigma_{1'}$	H2'-H6/8 ^a	H3'-H6/8 ^a	Ratio	%S
G1	-	-	-	2,61	4,53	0,58	97
C2	8	5,2	13	2,44	3,83	0,64	90
C3	4,5	8	11,4	3,03	3,2	0,95	44
G4*	1,9	7,1	10,4	3,74	2,55	1,47	<0
G5*	8,8	4,4	13,5	2,9	4,09	0,71	82
G6	-	-	-	2,8	4,85	0,58	96
T7	9,5	4,9	14,5	2,46	3,84	0,64	89
C8	9,7	5,2	14,9	1,83	3,05	0,6	94
G9	-	-	-	1,99	3,59	0,55	99
C10	-	-	-	2,2	3,3	0,67	86
Base	$J_{1'2'}$	$J_{1'2''}$	$\Sigma_{1'}$	H2'-H6/8 ^a	H3'-H6/8 ^a	Ratio	%S
G11	-	-	-	2,39	4,01	0,6	94
C12	10	5	16,3	1,86	4,25	0,44	>100
G13	-	-	-	2,22	4,5	0,49	>100
A14	8,8	5,3	14	2,66	3,49	0,76	76
C15	4	7,7	12,4	2,9	3,34	0,87	64
C16	9,1	4,6	13,7	2,52	2,7	0,93	57
C17	11	4,4	15,8	2,45	4,8	0,51	>100
G18	-	-	-	2,68	3,8	0,71	82
G19	-	-	-	2,73	4,28	0,64	90
C20	-	-	-	-	3,42	-	-

TAB. 2.3 – Couplages spin-spin $J_{1'2'}$, $J_{1'2''}$ et $\Sigma_{1'}$ déterminés à partir des spectres DQF-COSY, distances inter-protons H2'-H6/8 et H3'-H6/8 déterminées à partir des spectres NOESY, et ratio $\frac{H2'-H6/8}{H3'-H6/8}$ pour le duplex platiné G*G*G. Une estimation de la conformation du sucre, calculée à partir du ratio $\frac{H2'-H6/8}{H3'-H6/8}$, est présentée dans la dernière colonne (voir le texte pour les explications du calcul). ^aLes distances inter protons H2'-H6/8 et H3'-H6/8 ont été moyennées à partir des trois spectres NOESY : temps de mélange de 200 ms à 305 K en D₂O, temps de mélange de 80 ms à 305K en D₂O et temps de mélange de 200 ms à 281 K en H₂O. Toutes les distances n'ont pas pu être mesurées pour chaque temps de mélange.

2.4.3 Indications sur les sucres à partir des expériences NOESY

Les distances intranucléotidiques H2'-H6/8 et H3'-H6/8 sont corrélées avec l'angle de phase P et l'angle glycosidique χ [204]. Comme aucun sucre *syn* n'a été détecté lors de l'analyse des sucres de G*G*G, ces deux distances peuvent être utilisées pour estimer la conformation des sucres [160, 204]. Nous avons combiné trois expériences NOESY pour déterminer les distances moyennes H2'-H6/8 ($d_{H2'app}$) et H3'-H6/8 ($d_{H3'app}$) à partir de l'intégration des volumes NOE, le volume de la distance cytidine H5-H6 (2,45 Å) servant de référence (voir tableau 2.3). Le ratio $\frac{d_{H2'app}}{d_{H3'app}}$ a permis de calculer la proportion de S en supposant un équilibre entre les populations N et S.

Dans une fraction contenant une proportion x_S de la phase S et $(1-x_S)$ de la phase N d'un oligonucléotide donné, la distance apparente $d_{H2'-H6/8}$ est proportionnelle, en absence de diffusion de spin, à $[x_S * d_{H2'S}^{-6} + (1-x_S) * d_{H2'N}^{-6}]^{-\frac{1}{6}}$, où $d_{H2'S}$ et $d_{H2'N}$ sont respectivement les distances H2'-H6/8 en conformation S ou N. De la même manière la distance apparente $d_{H3'app}$ est proportionnelle à $[x_S * d_{H3'S}^{-6} + (1-x_S) * d_{H3'N}^{-6}]^{-\frac{1}{6}}$, où $d_{H3'S}$ et $d_{H3'N}$ sont respectivement les distances H3'-H6/8 en conformation S ou N. Le rapport $\frac{d_{H2'app}}{d_{H3'app}}$ est défini dans l'équation 2.1 :

$$\frac{d_{H2'app}}{d_{H3'app}} = \left(\frac{x_S * d_{H2'S}^{-6} + (1-x_S) * d_{H2'N}^{-6}}{x_S * d_{H3'S}^{-6} + (1-x_S) * d_{H3'N}^{-6}} \right)^{-\frac{1}{6}} \quad (2.1)$$

Les valeurs moyennes sont issues des données calculées à partir des modèles minimisés extraits de simulations de dynamique moléculaire avec AMBER 6 et parm 98 (voir partie I et [185]). Les distances obtenues sont $d_{H2'S} = 2,45 \pm 0,32$ Å, $d_{H2'N} = 3,27 \pm 0,24$ Å, $d_{H3'S} = 4,49 \pm 0,26$ Å et $d_{H3'N} = 2,27 \pm 0,26$ Å. En utilisant ces valeurs et l'équation 2.1 nous avons tracé le ratio $\frac{d_{H2'app}}{d_{H3'app}}$ en fonction de x_S . Il apparaît que la courbe de régression associée est quasiment linéaire, le calcul peut être approximé comme le présente l'équation 2.2 :

$$\frac{d_{H2'app}}{d_{H3'app}} = \frac{3,27}{2,27} + x_S * \left(\frac{2,45}{4,49} - \frac{3,27}{2,27} \right) \simeq 1,44 - 0,89x_S \quad (2.2)$$

Nous avons utilisé l'équation 2.2 pour calculer les pourcentages de la phase S pour tous les nucléotides de dsG*G*G, ces pourcentages sont présentés dans le tableau 2.3. Trois valeurs sont supérieures à 100% et une inférieure à 0%. Ces valeurs calculées sont cependant bien en-dessous de l'incertitude expérimentale sur la conformation, estimée à 20%.

La plupart des nucléotides a un ratio $\frac{d_{H2'-H6/8}}{d_{H3'-H6/8}}$ de 0,7 ou moins, indicateur d'une phase majoritairement S. G*4 d'un autre côté a un ratio de 1,465 typique d'une phase N. Des valeurs intermédiaires sont obtenues pour C3, C15 et C16. Les expériences DQF-COSY avaient déjà indiqué que C3 et C15 présentaient un équilibre $N \rightleftharpoons S$ tandis que C16 avait des constantes de couplage $J_{1'2'}$ et $J_{1'2''}$ typiques d'un sucre S (tableau 2.3). Cette incohérence peut trouver son origine, en principe, dans les marges d'erreurs des expériences DQF-COSY et NOESY. La simulation de dynamique moléculaire pour G*G*A a indiqué que C16 présentait de fréquentes transitions entre les conformations N et S tandis que pour G*G*G (voir plus loin) la conformation est très majoritairement N. Ces données suggèrent que C16, la cytosine complémentaire de la base platinée en 3', présente un équilibre $N \rightleftharpoons S$ déplacé vers la conformation N en raison de la coordination du cisplatine.

2.5 Distances RMN

L'analyse détaillée du spectre RMN a permis d'extraire 224 distances. Ces distances seront comparées dans le chapitre suivant avec la simulation de dynamique moléculaire.

2.6 Réarrangement de l'adduit

La purification par HPLC a permis d'isoler l'adduit G4*G5*G6 (**1**) de l'adduit G4G5*G6* (**2**). Les expériences RMN réalisées à la sortie de la purification ont permis de déterminer que la proportion du premier adduit était au début de l'enregistrement des spectres RMN d'environ 90 %.

Le tableau 2.4 présente les déplacements chimiques observés pour l'adduit migré.

	H8/H6	H2/H5/CH3	H1'	H2'	H2''	H3'	H4'	H1/H3	H41	H42
G1	8,00	-	6,02	2,70	2,79	-	-	-	-	-
C2	7,55	5,44	6,10	2,20	2,50	4,89	-	-	8,42	6,53
C3	7,46	5,62	5,58	2,02	2,32	4,84	-	-	8,61	6,84
G4	7,86	-	6,03	2,56	2,76	4,96	-	13,15	-	-
G5*	8,29	-	6,02	2,19	2,75	4,78	-	-	-	-
G6*	7,95	-	5,64	2,24	2,57	-	-	13,16	-	-
T7	7,50	1,32	6,17	2,25	2,56	4,91	4,27	13,97	-	-
C8	7,50	5,68	5,77	-	-	-	-	-	8,58	6,91
G9	-	-	-	-	-	-	-	13,09	-	-
C10	-	-	-	-	-	-	-	-	-	-
	H8/H6	H2/H5/CH3	H1'	H2'	H2''	H3'	H4'	H1/H3	H41	H42
G11	-	-	-	2,28	-	-	-	-	-	-
C12	7,42	5,42	5,77	-	-	-	-	-	-	6,52
G13	7,94	-	5,72	-	-	-	-	12,94	-	-
A14	8,18	7,90	6,29	2,68	2,86	-	-	-	-	-
C15	7,46	5,53	5,95	2,00	2,39	4,79	-	-	8,41	6,86
C16	7,57	5,49	5,98	2,11	2,45	4,80	-	-	8,20	6,70
C17	7,39	5,50	5,63	2,00	2,36	4,84	-	-	8,43	6,82
G18	7,87	-	5,64	2,70	-	5,00	-	13,15	-	-
G19	7,81	-	6,03	2,59	2,77	5,00	-	-	-	-
C20	7,46	5,47	6,22	-	-	-	-	-	-	-

TAB. 2.4 – Déplacements chimiques pour ^1H du duplex GG*G*. Tous les spectres ont été enregistrés à 305 K à l'exception des spectres pour les protons échangeables qui ont été enregistrés à 281 K. Les déplacements ont pour référence le TSP (Trisilyl Méthyl Phosphate).

Au cours de l'analyse RMN, il est apparu que l'adduit G*G*G était en train d'évoluer (figure 2.5). Une faible proportion de la nouvelle espèce (environ 10 %) était

présente dès le début de l'analyse, et cette proportion a rapidement évolué en 180 jours pour devenir majoritaire (la proportion entre nouvelle espèce et l'ancienne s'est alors stabilisée à 55 / 45).

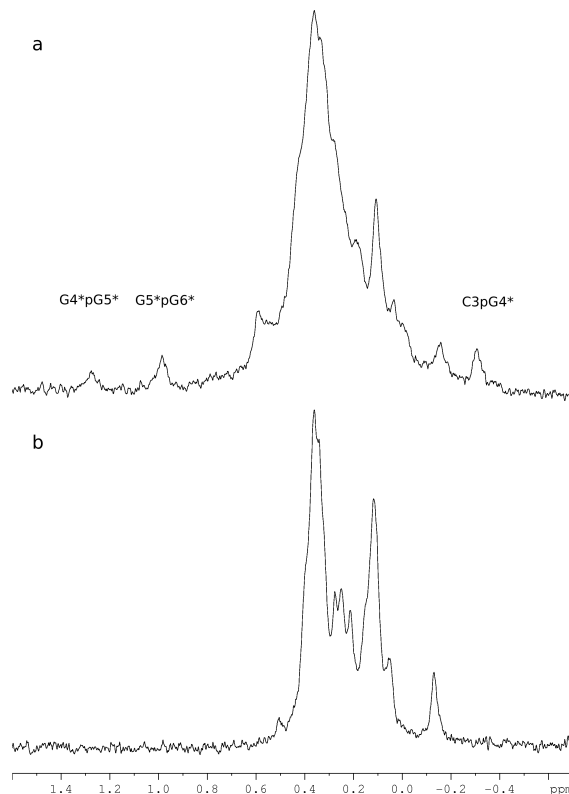


FIG. 2.5 – Déplacements chimiques du phosphore où la présence de G^*G^*G et GG^*G^* est détectée. (a) spectre enregistré 100 jours après la platination, les deux adduits sont présents. (b) spectre de l'oligonucléotide non platiné. Les déplacements chimiques sont présentés avec pour référence l'acide phosphorique à 85 %, l'enregistrement a été effectué à 305 K.

Après analyse de la nouvelle espèce, nous avons déterminé que le produit de réarrangement est le chélate GG^*G^* (**2**) (dont le profil apparaît en HPLC, voir figure 1.1). Le diagramme du réarrangement de (**1**) est présenté figure 2.6, sur la base de la mesure des pics de corrélation entre les protons H5-H6 des cytidines. Le premier point expérimental à 67,6% à 9 jours fait l'objet de vérifications complémentaires. Le temps représenté indique le temps où l'échantillon a été repris en solution, à une température normalisée pour les différentes analyses à 283 K. Cette renormalisation permet de prendre en compte le fait que l'augmentation de température de 10 K double le taux d'interconversion des deux espèces. Le temps t_{eff} qui résulte de cette normalisation est calculé suivant l'équation 2.3 :

$$t_{eff} = 2t \frac{(T - 283)}{10} \quad (2.3)$$

t est le temps réel passé à la température T .

En supposant une réaction d'ordre 1 pour la conversion de l'espèce A (G^*G^*G) en l'espèce B (GG^*G^*), l'équation différentielle est de la forme 2.4 :

$$\frac{dA}{dt} = -k_f A + k_b B \quad (2.4)$$

k_f et k_b sont les constantes de conversion de A vers B et de B vers A.

Si à $t=0$, la concentration de A est A_0 et celle de B est zéro, nous obtenons A en fonction du temps :

$$A = A_0 - \frac{1}{(k_f + k_b)} [k_b + 1e^{(-k_f+k_b)t}] \quad (2.5)$$

La régression résultant de l'équation 2.5 des points expérimentaux est présenté sur la figure 2.6. En excluant le point à 9 jours qui fait l'objet de vérifications, cette régression est en parfait accord ($R=0,99$) avec les observations, ce qui conforte l'hypothèse d'un réarrangement de cinétique d'ordre 1.

Cinétique du réarrangement pour ds G^*G^*G

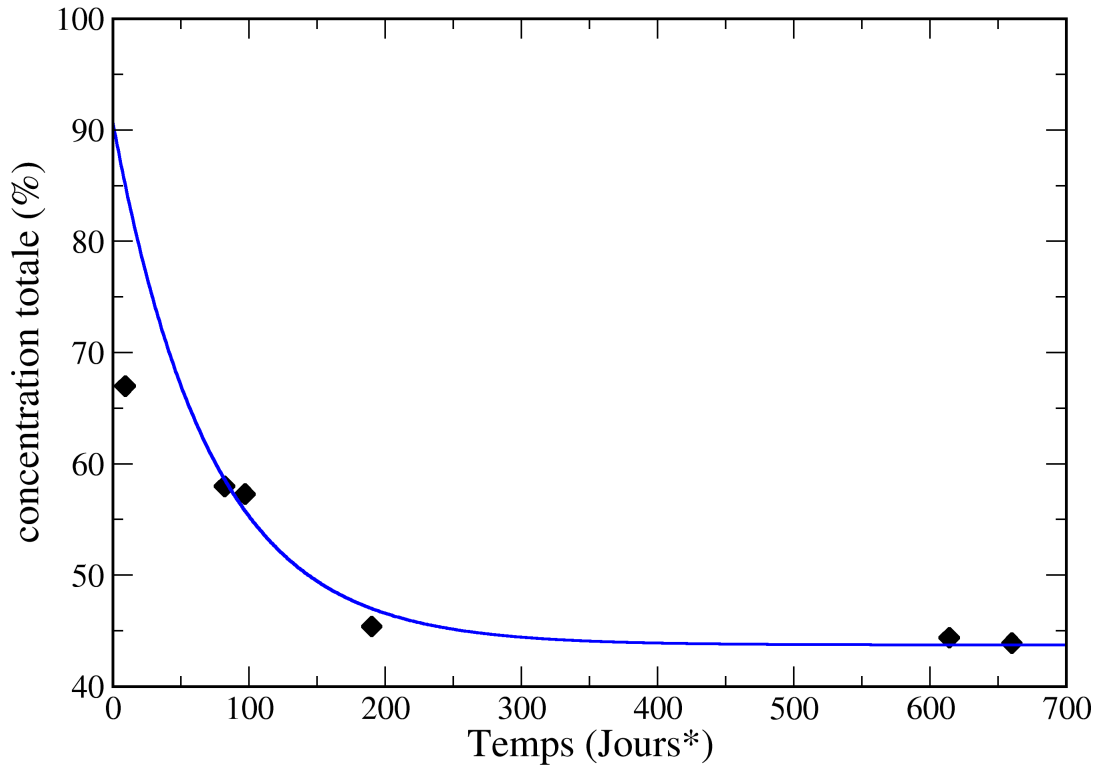


FIG. 2.6 – diagramme cinétique du réarrangement de G^*G^*G vers GG^*G^* . Jours* : le temps présenté est t_{eff} , le calcul de ce temps est décrit dans le texte.

Chapitre 3

Dynamique moléculaire

La simulation réalisée par le Dr. M.-A. Elizondo-Riojas [59] sur la séquence 5' - GCCG*G*ATCGC - 3' / 5' - GCGATCCGGC - 3' (notée G*G*A) a permis de décrire le comportement dynamique de l'ADN coordonné par le cisplatine. La corrélation a posteriori avec l'étude RMN effectuée par le Dr. F. Herman de cet oligonucléotide platiné [87] a donc permis d'affiner l'étude conformationnelle de l'adduit. En effet, les déplacements chimiques du proton H2' de C₃ et du proton H2' de C₁₇ ont été expliqués en dynamique moléculaire par le mouvement de va-et-vient horizontal de la paire de base G*₄-C₁₇ observé en dynamique moléculaire. L'analyse de la dynamique de l'adduit a révélé, du côté 3' du site de platination, la présence d'une conformation BII sur le phosphate entre C₆ et T₇ (noté C₆P₇ par la suite). Cette conformation, en proportion faible dans un ADN B [67], apparaît de manière concomitante avec la formation d'une liaison hydrogène entre l'ammine 3' du cisplatine et l'atome N₇ de l'adénine 6.

Lors de la simulation du docteur Dr. M.-A. Elizondo-Riojas, la conformation possédant la sous-population BII pour le pas A₆P₇ (ayant $\varepsilon - \zeta \geq 0$) est la plus importante. Cette sous-population représente 43% de la totalité de la simulation. Pour valider de manière expérimentale la présence de cette sous-population et en cherchant à l'augmenter, la séquence 5' - GCCG*G*GTCGC - 3' / 5' - GCGACCCGGC - 3' (G* représente une guanine platinée) a été choisie pour l'analyse par spectroscopie RMN. D'après les simulations préliminaires que j'ai effectuées, le remplacement de l'adénine A₆ par la guanine G₆ permet d'obtenir une augmentation de la proportion de la sous-conformation de 43% à 50%¹. En effet, l'atome N₇ de la guanine est plus électronégatif que celui de l'adénine [156] et l'atome O₆ apporte une possibilité supplémentaire de liaison hydrogène.

¹Ce pourcentage est donné à titre indicatif. En effet la correspondance de temps entre la simulation et les expériences RMN ne peut être absolue. De plus cette simulation a mis en évidence des limitations dans la paramétrisation du cisplatine, ce qui sera présenté dans ce chapitre page 90 et plus en détail dans la partie V.

3.1 Préparation de la dynamique

Le point de départ de la simulation a été le même que celui utilisé pour la simulation de G*G*A. Les bases A6 et T15 ont été remplacées en utilisant *xleap* par G6 et C15. Cette procédure a permis la création de la séquence $5' - \text{GCCG}^*\text{G}^*\text{GTCGC} - 3' / 5' - \text{GCGACCCGGC} - 3'$ (notée GGG pour le non platiné et G*G*G pour le platiné).

La structure ainsi obtenue a été minimisée dans le logiciel *xleap* de la suite AMBER 6.0 avec le champ de force parm 98 modifié pour prendre en compte le composé cisplatine (voir partie V). Ce champ de force modifié a été utilisé par la suite pour l'ensemble de la simulation. La préparation et la simulation ont été réalisées selon le protocole listé ci-dessous :

1. inclusion de 16 contre-ions ;
2. hydratation de l'ensemble ADN-cisplatine + contre ions dans une boîte orthorhombique en utilisant des molécules d'eau de type TIP3P. Les dimensions ont été définies pour avoir au moins 10,5 Å entre le bord de la boîte et le soluté (ADN+cisplatine+contre-ions), soit une taille finale de 64*54*54 Å.
3. minimisation pour ordonner les molécules d'eau et obtenir une structure stable pour le début de la dynamique (quand on crée la boîte d'eau, par défaut toutes les molécules d'eau sont positionnées selon un arrangement géométrique régulier et les molécules d'eau ne sont pas orientées correctement vis-à-vis du soluté) ;
4. chauffage ;
5. équilibration ;
6. production.

Ces étapes et les fichiers associés sont détaillés dans la partie V.

L'ensemble de la simulation a été calculée sans contrainte et seules les dix dernières nanosecondes de simulation de dynamique moléculaire ont été retenues pour l'analyse et la comparaison avec les données expérimentales issues de l'expérience de RMN.

3.2 Analyse de la dynamique

L'ensemble des paramètres hélicoïdaux calculés par Curves [112, 113] et mis en forme par l'intermédiaire du script analyse.pl (page 288) sera présenté en *Supplementary Material* dans l'article définitif (23 pages). En effet, comme nous le verrons ci-dessous la simulation ne permet pas de reproduire complètement les expériences RMN. Un travail est en cours pour prendre en compte les déformations spécifiques induites par le cisplatine dans le champ de force d'AMBER (voir partie V). Les points sui-

vants présenteront les caractéristiques principales de la coordination du cisplatine sur la séquence G*G*G.

3.2.1 Caractéristiques générales issues de la simulation

L'ADN simulé présente un coude de $67 \pm 9^\circ$ essentiellement localisé au niveau du site de platination (roulis local *-roll-* de $74 \pm 3^\circ$ entre G₄* et G₅*). Le mouvement dynamique de va-et-vient observé avec la paire G₄:C₁₇ pour G*G*A est fortement diminué dans le cas de G*G*G : ces observations sont corroborées par le blindage plus important dans le cas de G*G*G sur C₃H_{2'} (1,52 ppm au lieu de 1,58 ppm) et plus faible sur C₁₇H_{2'} (1,98 ppm au lieu de 1,96 ppm).

3.2.2 Conformation des sucres

Le tableau 3.1 présente la comparaison des conformations des sucres entre l'analyse (tableau 2.3) de la RMN et les calculs de dynamique moléculaire.

Base	%S _{RMN}	%S _{dyn}	Base	%S _{RMN}	%S _{dyn}
G1	97	100	G11	94	100
C2	90	100	C12	>100	99
C3	44	22	G13	>100	99
G4*	<0	21	A14	76	90
G5*	82	100	C15	64	66
G6	96	100	C16	57	5
T7	89	100	C17	>100	97
C8	94	100	G18	82	100
G9	99	99	G19	90	100
C10	86	96	C20	-	97

TAB. 3.1 – Comparaison entre les conformations des sucres observées en RMN et celles calculées à partir de la simulation de dynamique moléculaire.

Le champ de force parm 98 est connu pour déplacer l'équilibre N \rightleftharpoons S vers la conformation S [29]. Il apparaît de la comparaison issue du tableau que cette surreprésentation de S peut aller jusqu'à 20% (G₄*, G₅*, G₁₈). De plus, la coordination du cisplatine sur son ADN cible induit une orientation vers N de la guanine 5' platinée [151]. Les phases des sucres de C3 et C16, en tenant compte de ces remarques, présentent donc une incohérence entre les observations expérimentales et la simulation de dynamique moléculaire.

3.2.3 Conformation des phosphates

Lors de la simulation du Dr. Elizondo-Riojas, il est apparu qu'en 3' de G*G*A (sur ApT), la conformation BII était l'espèce majoritaire et couplée à la présence d'une liaison hydrogène entre G₆:N₇ et l'amine 3' du cisplatine. Un des objectifs de cette étude était l'augmentation de la proportion de BII pour permettre son observation en RMN. Selon l'analyse du spectre du phosphore en fonction du temps (figure 2.5), seuls les phosphates entre les guanines platinées présentent un signal particulier, ce qui est conforme avec les autres expériences RMN concernant le cisplatine [4]. L'approche développée dans la partie I n'a pas non plus indiqué de conformation BII prépondérante pour ce phosphate.

Durant la dynamique, le pas G₆pT₇ est entièrement BI, en accord avec les observations RMN et le calcul alternatif utilisant les combinaisons de distances inter-protons [185]. La liaison hydrogène est néanmoins présente durant 75% du temps de la dynamique (G₆:N₇/ G₆:O₆ - ammine 3' du cisplatine).

Lors de l'analyse plus précise de la simulation de dynamique moléculaire, il apparaît qu'une transition $\alpha/\gamma : g^+/g^- \rightarrow \alpha/\gamma : t/t$ est présente entre G₆pT₇, ce qui défavorise la conformation BII [192].

3.2.4 Comparaison des distances RMN avec le modèle dynamique

L'analyse des NOEs avec les 221 distances correspondantes dans la dynamique moléculaire a indiqué des incohérences ciblées sur quelques pas :

- 186 NOEs sont en accord avec les distances calculées pour les 10 ns de la simulation
- 35 NOEs sont en désaccord (dont 16 aux extrémités de l'oligonucléotide, dans les deux dernières paires de bases)
- les 19 NOEs qui restent en désaccord au coeur de l'oligonucléotide sont reliés aux problèmes de proportion des sucres présentés dans le tableau 3.1.

3.2.5 Validation du modèle simulé par les données expérimentales

Comme observé dans les conformations des sucres, des phosphates et des 221 NOEs qui ont pu être isolés et validés, la simulation de dynamique moléculaire a permis d'élucider la structure tridimensionnelle de l'adduit platiné sur la séquence d(GCCG*G*GTCGC)-(GCGACCCGGC). Les différences observées au niveau des sucres sont issues de problèmes connus en simulation [29] mais relèvent aussi des limitations

de notre modèle de paramétrisation du cisplatine (le travail est en cours pour améliorer cette représentation, voir partie V). La conformation $\alpha/\gamma : t/t$ anormale entre G_6 et T_7 induit un déroulement particulier qui peut être responsable des variations observées sur la phase de C_3 et C_{16} . En effet, en perturbant le comportement naturel du squelette phosphodiester, cette conformation induit une rigidité qui répercute les déformations issues de la coordination du cisplatine sur C_3 et C_{16} .

La figure 3.1 présente la structure moyenne minimisée de l'adduit platiné.

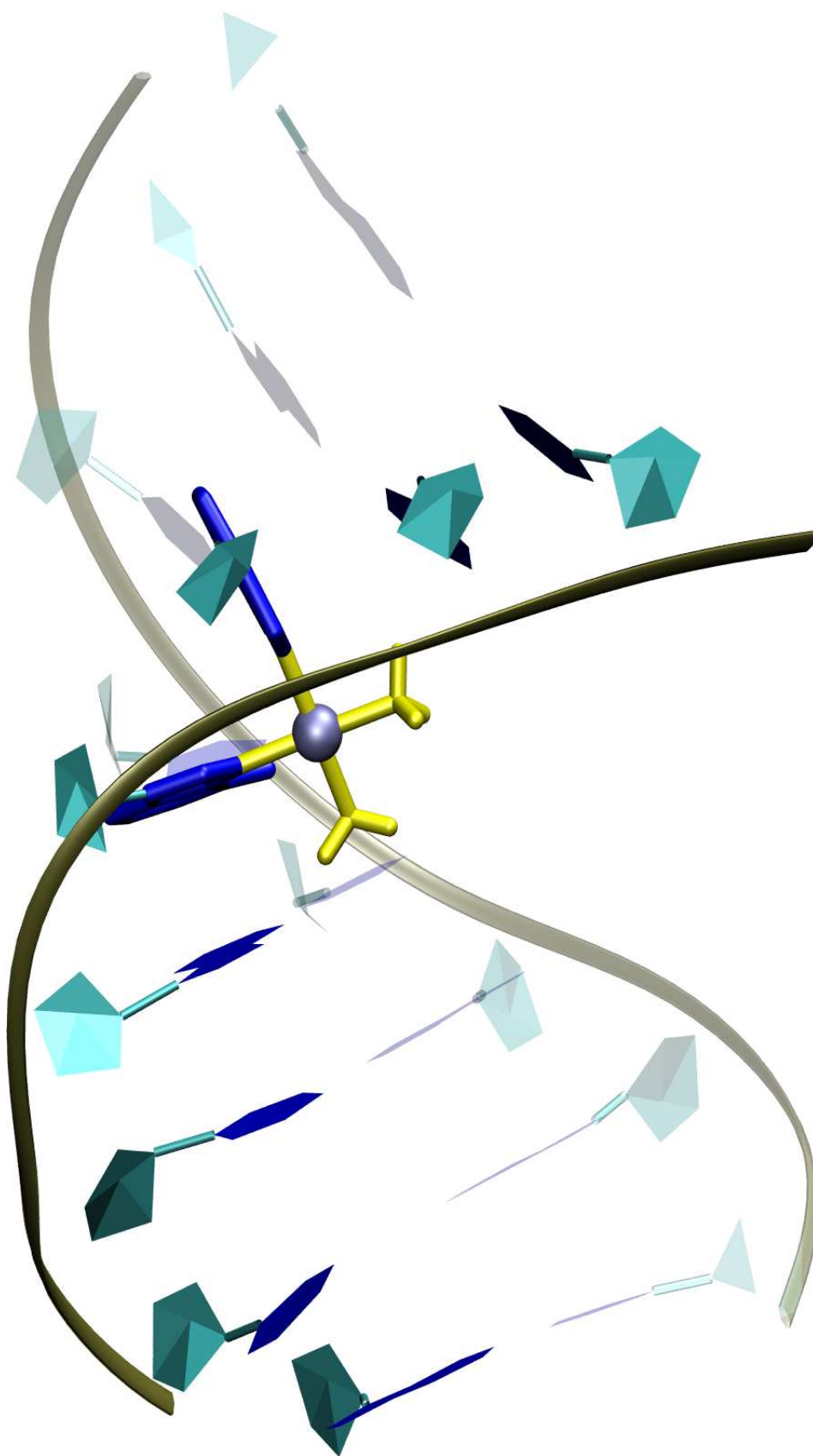


FIG. 3.1 – Représentation *NewRibbon* et *Licorice* [95] de la structure moyennée et minimisée issue de la simulation de dynamique moléculaire. Les guanines platinées sont en bleu, les ammines du cisplatine en jaune, le platine en cyan.

Chapitre 4

Conclusion

Cette étude est la première à mettre en avant la structure d'un adduit ADN-cisplatine sur une séquence comportant trois guanines adjacentes. L'isolement et la purification en HPLC ont permis d'isoler les différents adduits de la platination et d'obtenir l'adduit G*G*G en concentration suffisante pour permettre l'analyse en RMN de l'adduit formé.

Avec le croisement des techniques classiques de RMN (NOESY, COSY, TOCSY, DQF-COSY, ...) et de la simulation de dynamique moléculaire, la structure tridimensionnelle de l'adduit a pu être résolue.

Il apparaît que la structure BII présumée par Elizondo-Riojas et Kozelka à la suite de la simulation sur G*G*A n'est pas présente en RMN. De plus, la structure G*G*G est plus rigide que celle de G*G*A. Cette rigidité provient de la transition A vers G et par la stabilisation de la structure par le renforcement des liaisons hydrogènes entre l'amine 3' de l'adduit platiné et la purine qui suit.

Lors de l'analyse il est apparu que le modèle de représentation des atomes *parm 98* et celui élaboré au laboratoire pour le cisplatine présentaient quelques lacunes sur les équilibres conformationnels des sucres et des phosphates. Un travail de correction de la représentation du cisplatine a été initié (voir partie V) mais n'est pas encore satisfaisant pour être inclus dans l'étude présentée dans ce chapitre. Les travaux de reparamétrisation sont actuellement poursuivis par le Dr Djuranovic et feront l'objet d'un article séparé.

L'analyse de l'échantillon ayant été étalée dans le temps (2 ans), il a été possible de suivre le réarrangement de l'adduit G*G*G en GG*G* et de déterminer une première estimation de la cinétique du réarrangement. L'étude cinétique plus complète est actuellement réalisée par K. Chvalova.

Troisième partie

Reconnaissance plausible d'un ADN platiné par la LEF-I

L'ADN est la cible principale du cisplatine. Une fois le cisplatine coordonné sur sa séquence préférentielle (GG, AG), c'est la reconnaissance de ce dommage qui va induire l'efficacité thérapeutique. La première étape consiste en la fixation d'une protéine de type HMG, qui va ensuite entraîner une cascade de réaction cellulaire conduisant soit à la l'excision de l'adduit et à la réparation du brin, soit à l'apoptose. Les points de contrôle des différentes étapes sont de mieux en mieux connus mais font encore l'objet de nombreuses recherches [68, 69].

Le domaine accessible à la simulation se cantonne encore à l'heure actuelle à l'étude de composés de taille réduite, de quelques dizaines à quelques centaines de milliers d'atomes. Il n'est donc pas encore possible de représenter de manière explicite tous les mécanismes qui interviennent dans l'apoptose ou la réparation cellulaire. Il est néanmoins possible de commencer à étudier la reconnaissance de l'adduit platiné par les protéines HMG (High Mobility Group).

La facteur de transcription LEF-I (Lymphoid Enhancer-binding Factor I) fait partie de la famille des HMGs. C'est une protéine qui reconnaît spécifiquement la séquence 5'–GAGCTTCAAAGGGTG–3' en insérant une méthionine dans le désempilement créé par le coude de reconnaissance au niveau des adénines 23 et 24. La structure résultant de cette déformation ressemble à celle provoquée par la chélation du cisplatine sur deux guanines. Les expériences d'affinité sur gel réalisées par le Dr. Drumm ont validé que la mutation de A₂₃ et A₂₄ pour G₂₃^{*} et G₂₄^{*} (ou G^{*} représente une guanine platinée par le cisplatine) permet de conserver l'affinité de reconnaissance entre la séquence cible et LEF-I. L'affinité mesurée est comparable pour les deux séquences.

Dans un premier temps, nous avons cherché à décrire la structure de la protéine fixée sur sa séquence native cible.

La simulation de l'ensemble ADN-protéine (non platiné) a permis de décrire l'incorporation de la méthionine 11 dans le désempilement issu de la fixation de la protéine. L'utilisation de conditions périodiques et de solvant explicite a indiqué que cette fixation s'accompagne de la possibilité pour une molécule d'eau de s'insérer dans la poche créée par le mouvement de la méthionine. Durant la simulation de 10 ns sans contrainte, cette poche hydrophile a accueilli une première molécule d'eau pendant 3,5 ns qui a par la suite été remplacée par une seconde.

Cette simulation a mis en évidence la possibilité de reproduire le modèle de reconnaissance d'une protéine sur son ADN cible, avec l'implication possible d'une molécule d'eau. De plus, les expériences sur gel ont montré que le remplacement des adénines pour des guanines platinées conserve l'affinité de la protéine pour l'ADN. Ces deux résultats suggèrent que le modèle de reconnaissance ADN-platiné/LEF-I peut adopter une conformation telle que celle observée dans la simulation présentée dans l'article.

Ce travail est donc une première étape plausible vers la compréhension de l'initiation

de la cascade réactionnelle qui mènera la cellule vers la réparation du dommage causé par le cisplatine ou vers l'apoptose.

For submission to *J. Biomol. Struct. Dyn.*

Recognition Complex between the HMG Domain of LEF-1 and its Cognate DNA Studied by Molecular Dynamics Simulations with Explicit Solvation

Markus Drumm,[#] Stéphane Téletchéa,[#] and Jiří Kozelka^{*}

Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques, UMR 8601
CNRS, Université René Descartes, 45, rue des Saints-Pères, 75270 Paris Cedex 06, France

[#] These two authors contributed equally to the present work.

^{*} Corresponding author. Tel: +331 42 86 20 86; Fax: +331 42 86 83 87; E-mail: jiri.kozelka@univ-paris5.fr

***Running title:* Molecular Dynamics of a DNA-LEF-1 Complex**

Abstract

Molecular dynamics simulations of the complex formed between the HMG box of the lymphoid enhancer-binding factor (LEF-1) and its cognate DNA duplex were carried out with explicit inclusion of water. The simulation started with an NMR-based model (pdb code 2LEF) and the dynamics was pursued for 10 nanoseconds without constraints. It revealed that water intervenes in many ionic/polar interactions, establishing in particular local equilibria between direct and water-mediated hydrogen bonds, and thus increasing the entropy of the complex. Quite unexpectedly, the simulation indicated that a binding pocket for a specific water molecule may be reversibly formed at the apex of the bend induced in the DNA helix by LEF-1 binding, where a methionine side chain intercalates between two destacked adenines. We observed that the specific water molecule can temporarily replace the intercalated S-CH₃ group, acting as a sort of "extension" of the side chain. The residence time of this water molecule was about 3.5 ns. Simulations of the cognate DNA alone showed that this sequence has no intrinsic tendency to bend; therefore, the bending occurs solely as a consequence of the recognition, following the "induced-fit" mechanism.

Introduction

HMG-box (HMGB) proteins bend DNA and bind preferentially to distorted DNA structures. Their recognition of specific DNA sites results from a combination of hydrophobic interactions between the "hydrophobic cores" of the HMG-box and the minor groove of DNA, electrostatic interactions between basic linker regions of the protein and the sugar-phosphate DNA backbone, and base-specific hydrogen bonds (1, 2). A characteristic feature of DNA recognition by HMGB proteins is intercalation of a hydrophobic protein residue into the cleft produced by destacking of two base-pairs at the bending site(s).

Several structures of HMGB protein complexes with DNA have been solved by NMR (3-5) or X-ray crystallography (6, 7). While these structures provide valuable structural information about the recognition "strategies" of HMGB proteins, they allow for only limited insights into the dynamics of the binding. Molecular dynamics (MD) simulations has recently become a reliable tool for studying the thermal movement of proteins and nucleic acids, especially since the particle-mesh-Ewald (PME) technique has allowed the problem of representing long-range electrostatic forces to be overcome (8).

In the present paper, we report molecular dynamics simulations of the HMG-box of a sequence-specific HMGB protein, the transcription factor LEF-1, complexed with its cognate 15 bp DNA, as well as of the DNA alone. We were interested in the LEF1-DNA complex because of the similarity between the DNA bend locus seen in this complex and DNA bent by cisplatin crosslinking (7). We have recently studied the dynamics of the cisplatin-DNA crosslink (9), and it seemed interesting to us to compare this case of a DNA bend induced by a covalent crosslink with that imposed by a non-covalent interaction with a protein.

All simulations took account of the solvent explicitly and used the particle-mesh-Ewald method (10) to determine electrostatic interactions. The recognition complex of LEF-1 with its DNA sequence has been previously studied by NMR (3), and the NMR coordinates were utilized as starting conformation for the simulation (Figure 1). We show here that the simulation of the free DNA converges towards the same linear structure, whether started from the DNA coordinates of the NMR study of the complex or from canonical B-DNA. Analysis of the MD trajectories indicated that the recognition DNA sequence of LEF-1 does not present a natural tendency for bending toward the major groove, consistent with an "induced-fit" mechanism (11, 12). We have used the analysis of our MD simulation to address questions such as those of the dynamics of the intercalation site, sugar pucker equilibria, the dynamics of hydrogen bonds, and the role of BI \rightleftharpoons BII interconversions.

The amino acid and DNA sequences are given in Figure 2 along with the respective numbering schemes. It has to be noted that in the first paper of Love et al. (3), the residue numbering for the HMG domain of LEF-1 started with 0 and is therefore shifted by -1 with respect to that used in their second paper (13) and to that adopted in this work.

Insert Figures 1 and 2

Methods

The MD simulations were carried out with the SANDER module of the program AMBER Version 6.0. (14) run on a cluster of IBM-compatible personal computers under Red Hat GNU/Linux 7.1, or on an IBM Power4 P690 computer operating under AIX 5.1. The parm98 forcefield (15) was used without modification. The particle-mesh-Ewald method (10, 16) using charge grid spacing of approximately 1 Å with cubic B-spline interpolation and sum tolerance of 10^{-5} Å was used to calculate the electrostatic energy. A 9 Å cutoff was applied to Lennard-Jones interactions. The MD simulations employed the SHAKE algorithm (tolerance=0.0005) to all X-H bonds (17) and used a time step of 2 fs.

The initial coordinates for the protein-DNA-complex were taken from the energy-minimized mean structure obtained upon averaging the 12 individual structures of the PDB-entry for the NMR-derived structure. Two simulations of the 15 bp cognate DNA were performed: one starting from the initial structure used for the simulation of the protein-DNA complex (vide supra) and another from canonical Arnott B-DNA (18) constructed using the NUCGEN module of the AMBER package.

In order to neutralize the systems, the appropriate number of sodium counter ions (28 for the free DNA, 16 for the protein-DNA-complex) was placed at the phosphate bisectors approximately 6 Å from the oxygen atoms. In the case of a close contact of a charged side chain amino-acid with the phosphate, no counter ion was added. The solute was then solvated in a rectangular box of randomly oriented water molecules (approx. dimensions: 75x74x64 Å for the protein-DNA complex and 80x60x50 Å for the free DNA). The TIP3P water model was employed for all simulations along with periodic boundary conditions.

The first rounds of energy minimization served to relax the water molecules while keeping fairly high positional restraints on the solute atoms. Only the final round of energy minimization was done unrestrained (for the protocol, see ref. (9)). The system was then heated to 300 K in two rounds of 20 ps with harmonic constraints of 25 Kcal/mol.Å² imposed on the solute and counter ions in NVT conditions. During the first round, (0-150 K), the velocities were rescaled, whereas during the second round, the system was coupled to a heat bath using the Berendsen algorithm (coupling constant of 0.2 ps). The system was subsequently relaxed in six rounds of 20 ps each during which the harmonic constraints were subsequently relaxed to 5 Kcal/mol.Å² for the solute and 2.5 Kcal/mol.Å² for the counter ions. The seventh round was 100 ps with harmonic constraints of 2.5 Kcal/mol.Å² for the solute only. At this point, a Maxwell-Boltzmann redistribution of velocities was carried out. Four more rounds of equilibration were carried out (3x100, 1x200 ps) with position constraints of 2.5, 1.0, 0.5 and 0.1 Kcal/mol.Å² respectively applied to the hydrophobic core of the protein. Finally, 250 ps under the production conditions were run before the actual production was started. The center of mass velocity was removed every 100 steps. In order to avoid occasional openings of the terminal base-pairs, their Watson-Crick hydrogen bonds were reinforced by means of soft distance constraints of 10 Kcal/mol.Å². A simplified equilibration protocol, corresponding to that used previously in ref.(9), was employed for the simulations of the free DNA.

The atomic coordinates were stored every 1 ps. The analysis used the program CURVES(19) for calculation of helical parameters and backbone torsion angles and the

CARNAL and PTRAJ modules of AMBER 6.0 to monitor other structural features such as distances, angles and torsions.

Results and Discussion

Global agreement between the MD simulation and the NMR-derived static models

The initial structure used to start the MD simulations was constructed by averaging the 12 published NMR-based models (3) and minimizing the energy for 20 steps of steepest descent and 1053 steps of conjugate gradient minimization. The heated and equilibrated protein-DNA complex was followed for a production period of 10 ns without any constraints. The rms deviation from the initial structure rose to 2 Å during the first 1 ns and then stayed between 2 and 2.7 Å for the remaining time (not shown). As shown in Figure 3a, the main structural feature, i.e., the directional bend of the DNA double helix towards the major groove at the recognition site, which is common to all the 12 NMR models published earlier (3), is also observed during the simulation, as is the enlarged minor groove at the central base-pairs C5-G26 to T8-A23 (10-11.5 Å as measured by CURVES (19), not shown). Thus, the major structural aspects of the DNA at the recognition site as observed in the NMR models, are conserved during the simulation.

Insert Figure 3

In order to assess the extent of local conformational changes within the DNA duplex, we have plotted the torsion angles along the backbone and the principal helicoidal parameters as a function of time, and compared them to the values of the 12 NMR structures deposited in the NDB. During the whole simulation period, the torsion angles α , β , and γ remained in their canonical ranges (i.e., α , β^t , γ^+), except for a fortuitous excursion of C5 to $\alpha^t\gamma^t$. The torsion angle δ , ϵ , and ζ showed some fluctuations characteristic of N-S conformational transitions of sugars and BI-BII transitions of phosphate residues (shown for the central five base-pairs in Figure 4). This comparison revealed one principal deviation of the simulated structure from the NMR models: a BI \rightarrow B0 (or $\epsilon^t\zeta^- \rightarrow \epsilon^t\zeta^t(20, 21)$) transition at the G26/G27 step. This transition occurred during the heating period, and became manifest in an increased twist at the step C4:G27/C5:G26, and a decreased twist at the step C5:G26/T6:A25, as compared to the NMR models (Figure 5). This BI \rightarrow B0 transition could either be metastable, if the simulated structure was "trapped" in the B0 conformation during the heating period, or free-energy driven, which would mean that the previous refinement using NMR constraints

did not converge towards the correct free-energy minimum. As a check, we have switched the BI conformation of the G26/G27 step to B0 using a constraint on the ζ dihedral angle of nucleotide G26, and continued the simulation for further 5 ns, after which the constraint was removed and the structure relaxed for another 5 ns. As can be seen in Figures 3 and 4, the B0 conformation switched spontaneously back to BI as soon as the constraint was removed, and remained BI until the end of simulation. This suggests that the parm98 forcefield in combination with the TIP3P water model favors the BI conformation at the G26/G27 step. The ultimate test would consist in comparing the original NMR data with the structural features of our dynamic model; this was unfortunately not possible since the NMR data were neither deposited nor available from the authors.

Insert Figures 4-7

Dynamic water bridges: direct versus water mediated hydrogen-bonding interactions

The potential energy of the solute (excluding the water molecules) shows fluctuations up to 450 kcal/mol; these fluctuations are compensated by opposite fluctuations of the solvation energy (Figure 6). The fluctuations are caused mainly by interactions between ionic and/or hydrogen-bonding groups which undergo transitions between direct and water-mediated interactions. Most of the ion-pairs involved in such transitions involve lysines or arginines as cations, and glutamates or DNA phosphates as anions. The most notable fluctuations are shown in Figures 7a (protein-DNA) and 7b (protein-protein) as plots of characteristic interatomic distances as a function of time, together with the interatomic separations observed in Love's 12 NMR models (3). It can be seen that for many of the distances fluctuating during the MD simulation, there are also differences among the 12 NMR models (e.g., R46(C γ)-P8; N75(N δ)-P23). On the other hand, some of the distances predicted by the MD simulation to fluctuate are uniform among the NMR models (e.g., S30(O γ)-A25(N $_3$); K15(N ζ)-P25). Structural refinement using NMR constraints, such as applied in the study by Love et al. (3) cannot reproduce dynamic equilibria, and the constraints may drive the structure towards one of the energy minima, or towards an unphysical mean structure (22, 23). Therefore, our MD simulation is a valuable complement to the NMR study. It can not be excluded, however, that some of the discrepancies reflect insufficient statistical sampling in our simulation, so that our time average does not correspond to the NMR-observed time-average. Direct comparison between MD-simulated distances and the distance data used in the NMR refinement would resolve this ambiguity; unfortunately, the original distances were not available.

Protein binding modifies sugar conformations of the DNA

An important aspect of the adaptation of the DNA to the protein surface is the displacement of the N \rightleftharpoons S conformational equilibria of the deoxyribose rings. Structure refinement of these flexible parts using NMR-derived distance constraints can be problematic due to the r^{-6} dependence of the NOE volume (22). Unrestrained MD simulations, on the other hand, give a dynamic description of sugar pucker equilibria. The sugar puckers of the LEF1-DNA complex have not been discussed in the description of the NMR structure (3), however, the 12 models deposited in the pdb database do show some interesting particularities which can be compared with the results of our MD simulation. The phase angles P of the deoxyriboses are plotted for the central five base-pairs in Figure 4. Obviously, the trinucleotide C5:T6:T7 at the kink site is confined to N-type sugar puckers ("north", $0^\circ < P < 60^\circ$), as shown both by the NMR models and our MD simulation, although a certain flexibility of C5 is indicated by the simulation. The adjacent nucleotide, T8, strongly fluctuates between N and S ("south", $P > 120^\circ$) conformations, frequently assuming an E ("east", $60^\circ < P < 120^\circ$) conformation, and similar, although less dramatic, variations are observed between the different NMR models. The complementary nucleotide, A23, also shows fluctuations during the simulation, but with a more pronounced preference for either N or S, the S conformation being preponderant. In the NMR models, on the other hand, the A23 sugar pucker is exclusively N. This difference could be due to an artifact of the structure refinement under NMR constraints, since the parm94 forcefield of AMBER used is known to bias deoxyribose puckers towards the N conformation, whereas this tendency was counterbalanced (by increasing the V2 term for the O4'-C4'-C3'-O3' torsion angle) in the parm98 version (15) that was used in our MD simulations.

The repuckering of C5, T6 and T7 to N is quite obviously related to the local contraction of the backbone due to the bend induced on the DNA double helix. In the mean structure calculated over the first (unconstrained) 10 ns of the production period, the P-P-distances between the phosphates of C5pT6pT7 and T6pT7pT8 are shortened to 5.93 Å and 6.16 Å, respectively (canonical Arnott-B-DNA: 7.0 Å).

Protein binding displaces BI \rightleftharpoons BII conformational equilibria

Another “strategy” used by nucleic acids to match the surface of a recognition protein involves rotations of the phosphate residues about the ϵ and ζ torsion angles. Equilibria between the canonical $\epsilon'\zeta^-$ conformation (so-called BI conformation) and the alternative $\epsilon'\zeta^+$ (BII) or $\epsilon'\zeta^+$ (B0) conformations (20, 21) are observed in DNA at certain sequences, and shifts of these equilibria upon protein binding have been reported (21, 24). It can be seen from Figure 4 that for the homopyrimidine block C4-T8, the MD simulations and NMR models are in agreement and indicate that the phosphodiester groups are all fixed to the canonical BI conformation.

For the homopurine block of the opposite strand (A23-G27), the NMR models indicated that the nucleotide A25 is in BII and G26 in B0 conformation. In contrast, the MD simulation showed during the first 10 ns G26 in the canonical BI conformation whereas A25 was seen in a conformation closer to BII than B0 ($\epsilon \approx 210^\circ$, Figure 4). In addition, a fluctuating behavior for A23 was observed. As a check, we switched the conformation of G26 from BI to B0, constraining its ζ torsion angle to 180° ($k = 50$ Kcal/mol.rad²), and continued the simulation with the constraint for 5 ns. This induced an increase in the ϵ torsion angle of A25 towards gauche⁻, i.e., A25 approached the BII conformation seen in the NMR models. The base-pair twist angles at the steps C5:G26/T6:A25 and C4:G27/C5:G26 also moved towards values seen in the NMR models (Figure 5). When the constraint on G26 ζ was switched off (at 15 ns of the production period), the two twist angles returned to their original values, and the B0/BI conformation of the A25/G26 dinucleotide was restored. This alternative conformation of the backbone of the homopurine block identified by the MD simulation corresponds thus to a stable energy-minimum.

Insert Figures 8 and 9

Dynamics of the bend locus

Partial intercalation of hydrophobic residues between two destacked base-pairs has been identified as a principal “strategy” used by HMG-box (HMGB) proteins to bind and to bend DNA (1, 2), and has been observed so far in all molecular structures of HMGB-DNA complexes at atomic resolution (see Figure 3 in ref.(2)). In the case of LEF-1, the intercalating residue is methionine M11, and the intercalating site is situated between adenines A23 and A24 which are strongly destacked (note the large roll in Figure 5b). The intercalation of M11 seems to be assisted by M14, whose methyl group is in close contact to the one of M11.

The present MD study was expected, first, to bring a dynamic dimension to the static and time-averaged picture of site-specific recognition emanating from the NMR study, and second, to reveal an eventual role of specific water molecules in this recognition. In fact, our simulation indicated that the S-CH₃ group of methionine M11 may not remain fixed in the intercalation site but can exchange with another position further away from the cavity. Figure 7b shows the distances from the M11 methyl carbon (C_ε) to the N1 atoms of A23 and A24, and to the C_ε atom of the other methionine, M14, which remains practically stationary during the simulation. At the beginning of the simulation, the methyl carbon of M11 is at ~4 Å from A23N1, A24N1, and M14C_ε, whereas at ~1200 ps, the M11 methyl group leaves the cavity and all three distances increase. This displacement of the M11 methyl group is due partly to a rotation about the C_γ-C_δ bond, partly to displacements of the whole residue. After ~4600 ps, M11C_ε returns into the cavity and the short distances are restored. A detailed analysis has revealed that the site of M11C_ε in the intercalation site is taken over by a specific water molecule which is apparently crucial for the stabilization of the position of M11C_ε outside the gap between A23 and A24 (Figure 8). This water molecule ("H₂O A") occupies at the beginning of the simulation a cavity delimited by the amino acids M14, R18, S30, et I33, and the dinucleotide A24pA25 (Figure 8a). At 1075 ps, H₂O A leaves this cavity (Figure 8b) and when it passes between the methionines M11 and M14 (Figure 8b), it collides at 1272 ps with another water molecule ("H₂O B") which "pushes" it into the notch between A23 and A24, thus replacing the M11 S-CH₃ group. The latter is rotated and translated away from this intercalation site (Figures 8c+d). At ~3050 ps, H₂O B moves towards the cavity occupied initially by H₂O A (Figure 8e). Finally, at ~5000 ps, H₂O A leaves the intercalation site (Figures 8f+g). Figure 9 gives another account of the movements of the two water molecules, plotting their distances to the H2 atom of adenine A24 as a function of time. On perusal of Figures 5 and 7, it can be seen that the displacement of the M11 S-CH₃ group out of the intercalation site is accompanied by a diminution of the roll at the T7:A24/T8:A23 step.

Our MD simulation thus confirms, on one hand, that the intercalation of M11 between A23 and A24, as seen in all 12 NMR models, corresponds to a stable conformation; however, it suggests that an alternative conformation may exist where the intercalation site is taken over by a water molecule. The refinement of Love et al. could not have revealed such a conformation since the water molecules of the solvent were not explicitly modeled.

It is interesting to compare the dynamics of the bend locus between A23 and A24 with that of the previously studied bend induced by covalent crosslinking of two adjacent guanines by the antitumor drug cisplatin (9). In that case, the cavity between the destacked guanines is constantly filled with water molecules, and we suggested that solvation stabilizes the

destacking and increases the bend angle with respect to solid state structures. It is interesting to observe, in the present MD simulation, that water molecules can, in principle, help to fill the gap between destacked bases even in the case of a protein-DNA complex. Here the role of stabilizing the bend is basically fulfilled by the intercalating protein residue (M11) but our simulation suggests that the intercalation could be supplemented (or even alternated with) specific water molecules. The lifetime of these specific water molecules is indicated to be of several nanoseconds (Figure 9) and they should thus be detectable by NMR spectroscopy (25, 26).

Free DNA shows no intrinsic tendency to bend

The simplest recognition mechanism between proteins and substrates is the so called key and lock mechanism (12). If this mechanism applied to the recognition of the 15 bp DNA sequence by the HMG-domain of LEF-1, a sub-population of kinked DNA would be expected to exist even without the presence of protein. In order to test this hypothesis, we carried out two individual molecular dynamics simulations with the free 15 bp oligonucleotide. The first simulation was started with the DNA in the kinked conformation extracted from the complex with LEF-1. The second started with a standard B-DNA conformation. The bent DNA taken from the complex with LEF-1 straightened already during the equilibration procedure and the resulting MD production showed the same features as the one started from linear B-DNA. It is interesting to note that in a previously reported MD study of the structurally related complex between the HMG domain of human SRY protein and its cognate DNA (12), using CHARMM, the control simulation starting with the bent structure occurring in the complex, did not restore the linear B-DNA form of the free DNA previously determined by NMR (4).

Figure 3 compares the kink angle and the kink direction, as determined by the program CURVES (19), both for the free and the complexed DNA. In the complex a clear preference for high amplitude (40 to 90°) bending in the direction of the major groove (kink direction 180°) is observed. The free DNA shows only a tendency for slight (up to 45°), non-directional bending. We conclude that there exists no sub-population in the free DNA showing a directional kink towards the major groove, which renders the hypothesis of a pre-formed kink unlikely.

In line with the absence of predisposition for bending, we detected no tendency of the trinucleotide C5:T6:T7 to assume N puckers, or of the nucleotides A25 and G26 to assume B0 or BII conformations. Interestingly, in the free DNA, the phosphate A23pA24 is found almost

purely in BII conformation with only few transitions to BI, whereas in the complex, this equilibrium is clearly shifted towards BI.

The significant conformational changes that DNA undergoes upon forming the complex with LEF-1 do therefore not seem to be facilitated by intrinsic pre-disposition for such changes. Very recently, it has been shown by Wright et al. that the protein also undergoes significant restructuring upon complex binding (13). Thus, structural adaptation of both partners is necessary for the formation of the LEF1-DNA complex.

Conclusion

The present MD simulations of the solvated LEF1-DNA recognition complex and of the free DNA provided several interesting insights into the recognition mechanism, complementing the NMR studies by Love et al. (3, 13). First, the simulation of the DNA sequence alone showed that this sequence has no intrinsic tendency to bend towards the major groove. Therefore, the directional bend observed in the complex is a pure consequence of protein binding. Second, the simulation of the complex suggested some alternative local conformations to those found previously. For instance, the sugar conformation of A23 was indicated to be dynamic, with an important contribution of the S pucker, whereas all the NMR-based models converged to an N conformation. This may be related to the properties of the parm94 forcefield used in the NMR refinement, which biases deoxyribose sugars towards N. Another point where the simulations indicated a different local conformation was the phosphodiester group between nucleotides G26 and G27 which in the simulation had the canonical BI conformation whereas the NMR models showed a B0 conformation. Third, the explicit representation of water molecules made it possible to observe dynamic exchange between direct ionic and/or hydrogen-bonding interactions, and those mediated by one or more water molecules. Thus, the entropic cost of the general rigidification upon complexation may be partly compensated by fluctuations at the protein-DNA interface. Finally, we observed an interesting event where a specific water molecule replaced the S-CH₃ group of methionine M11 from the intercalation site between A23 and A24; this conformation persisted for nearly four nanoseconds.

Molecular dynamics simulations are a powerful tool for studying dynamic conformational equilibria of macromolecules, especially when they are backed up by NMR data. The value of the present study could have been considerably increased, if the original NMR data previously recorded for the same system were available. This underlines the

importance of depositing not only the atomic coordinates (which are biased by the technique of structural refinement) but also the original NMR constraints.

Acknowledgments

We are indebted to Dr. M.-A. Elizondo-Riojas for valuable technical assistance. M. D. was Marie-Curie postdoctoral fellow (Contract QLGA-CT200-52018) of the European Commission. We gratefully acknowledge financial support by the Association for International Cancer Research (Grant 00-321) and by COST (Action D20/003/00), and computer time from the IDRIS computer center of CNRS.

References

1. A. Travers, *Curr. Op. Struct. Biol.* 10, 102-109 (2000).
2. J. O. Thomas & A. A. Travers, *Trends Biochem. Sci.* 26, 167-174 (2001).
3. J. J. Love, X. Li, D. A. Case, K. Giese, R. Grosschedl & P. E. Wright, *Nature* 376, 791-795 (1995).
4. M. H. Werner, J. R. Huth, A. Gronenborn & G. M. Clore, *Cell* 81, 705-714 (1995).
5. F. H.-T. Allain, Y.-M. Yen, J. E. Masse, P. Schultze, T. Dieckmann, R. C. Johnson & J. Feigon, *EMBO J.* 18, 2563-2579 (1999).
6. F. V. Murphy, S. R. M. & M. E. A. Churchill, *EMBO J.* 18, 6610-6618 (1999).
7. U. M. Ohndorf, M. A. Rould, Q. He, C. O. Pabo & S. J. Lippard, *Nature* 399, 708-712 (1999).
8. T. E. I. Cheatham & P. A. Kollman, *Ann. Rev. Phys. Chem.* 51, 435-471 (2000).
9. M.-A. Elizondo-Riojas & J. Kozelka, *J. Mol. Biol.* 314, 1227-1243 (2001).
10. T. A. Darden, D. York & L. G. Pedersen, *J. Chem. Phys.* 98, 10089-10092 (1993).
11. K. Nadassy, S. J. Wodak & J. Janin, *Biochemistry* 38, 1999-2017 (1999).
12. Y. Tang & L. Nilsson, *Proteins Struct. Funct. Gen.* 31, 417-433 (1998).
13. J. J. Love, X. Li, J. Chung, J. Dyson & P. E. Wright, *Biochemistry* 43, 8725-8734 (2004).
14. D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, Jr., R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, R. J. Radmer, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. K. Weiner & P. A. Kollman, . (University of California, San Francisco, 1999).
15. T. E. Cheatham III, P. Cieplak & P. A. Kollman, *J. Biomol. Struct. Dyn.* 16, 845-862 (1999).
16. U. Essman, L. Perera, M. Berkowitz, T. A. Darden, H. Lee & L. G. Pedersen, *J. Chem. Phys.* 103, 8577-8593 (1995).
17. J. P. Ryckaert, G. Ciccotti & H. J. C. Berendsen, *J. Comput. Phys.* 23, 327-341 (1977).
18. S. Arnott, P. Campbell-Smith & P. Chandrasekharan, in *CRC Handbook of Biochemistry* . (1976), vol. 2, pp. 411-414.
19. R. Lavery & H. Sklenar, *J. Biomol. Struct. Dyn.* 6, 655-667 (1989).
20. B. Hartmann, D. Piazzola & R. Lavery, *Nucleic Acids Res.* 21, 561-568 (1993).
21. B. Hartmann, M. R. Sullivan & L. F. Harris, *Biopolymers* 65, 250-264 (2003).
22. D. A. Pearlman & P. A. Kollman, *J. Mol. Biol.* 220, 457-479 (1991).
23. J. M. Withka, S. Swaminathan, D. L. Beveridge & P. H. Bolton, *J. Am. Chem. Soc.* 113, 5041-5049 (1991).
24. D. Djuranovic & B. Hartmann, *J. Biomol. Struct. Dyn.* , 771-788 (2003).

25. J. W. R. Schwabe, *Curr. Op. Struct. Biol.* 7, 126-134 (1997).
26. M. Levitt & B. H. Park, *Structure* 1, 223-226 (1993).
27. A. Dalke, W. Humphrey & J. Ulrich, (VMD (Visual Molecular Dynamics). Theoretical Biophysics Group, University of Illinois and Beckman Institute, Urbana, USA, 1997).

Figure Captions

Figure 1: Schematic representation of the NMR structure of Love et al. (3) (energy-minimized average used as initial structure for the MD simulation). The DNA is represented in blue as a licorice model, with the two destacked adenines in dark-blue sticks, the protein is shown as a green ribbon, with the intercalating methionine M11 highlighted as red CPK model. The figure was generated using the programs VMD (27) (<http://www.ks.uiuc.edu/Research/vmd>) and POV-Ray (<http://www.povray.org>).

Figure 2. Sequences of the HMG-Domain of LEF-1 (a) and its cognate DNA (b). Helical domains of the protein are printed in bold and underlined.

Figure 3: Correlation between the kink angle (amplitude) and kink direction of a) the DNA in the complex with protein and b) of the free DNA without the protein during the molecular dynamics. The red squares indicate the parameters for the NMR structures.

Figure 4. Plot of the torsion angles ϵ (a) and ζ (b) of the sugar-phosphate backbone, and of the deoxyribose phase angles (c) of the central base-pairs, of the cognate DNA complexed with LEF-1, as a function of simulation time. Note that a dihedral constraint to G26 ζ (180° , $k = 50$ Kcal/mol.rad²) was applied between 10 and 15 ns. The values found in the 12 NMR models (pdb accession number 2LEF) are shown in the right panel.

Figure 5. Plot of the helicoidal parameters twist (a) and roll (b) (calculated using the "global" helix axis, as determined by CURVES (19)) of the central five base-pairs, of the cognate DNA complexed with LEF-1, as a function of simulation time. The values found in the 12 NMR models (pdb accession number 2LEF) are shown in the right panel.

Figure 6. Potential AMBER energy (black) and solvation energy (red; calculated using the Generalized Born method implemented in AMBER6.0) plotted as a function of simulation time.

Figure 7. Protein-DNA (a) and protein-protein (b) interatomic distances undergoing important fluctuations during the production period of the MD simulation. The values found in the 12 NMR models (pdb accession number 2LEF) are shown in the right panel.

Figure 8. Snapshots showing the dynamics of the intercalation site, as described in the text. Water molecules "A" and "B" are shown in red and silver, respectively. a) 1007 ps; b) 1075 ps; c) 1272 ps; d) 1274 ps; e) 3420 ps; f) 4587 ps; g) 4957 ps. The dashed lines connect the O atom of water molecule "A" with the H2 atom of adenine A23, and the C_ε atoms of M11 and M14 with each other.

Figure 9. Positions of the water molecules "A" (red) and "B" (silver) represented as the distance from the H2 atom of adenine A24, plotted as function of simulation time.

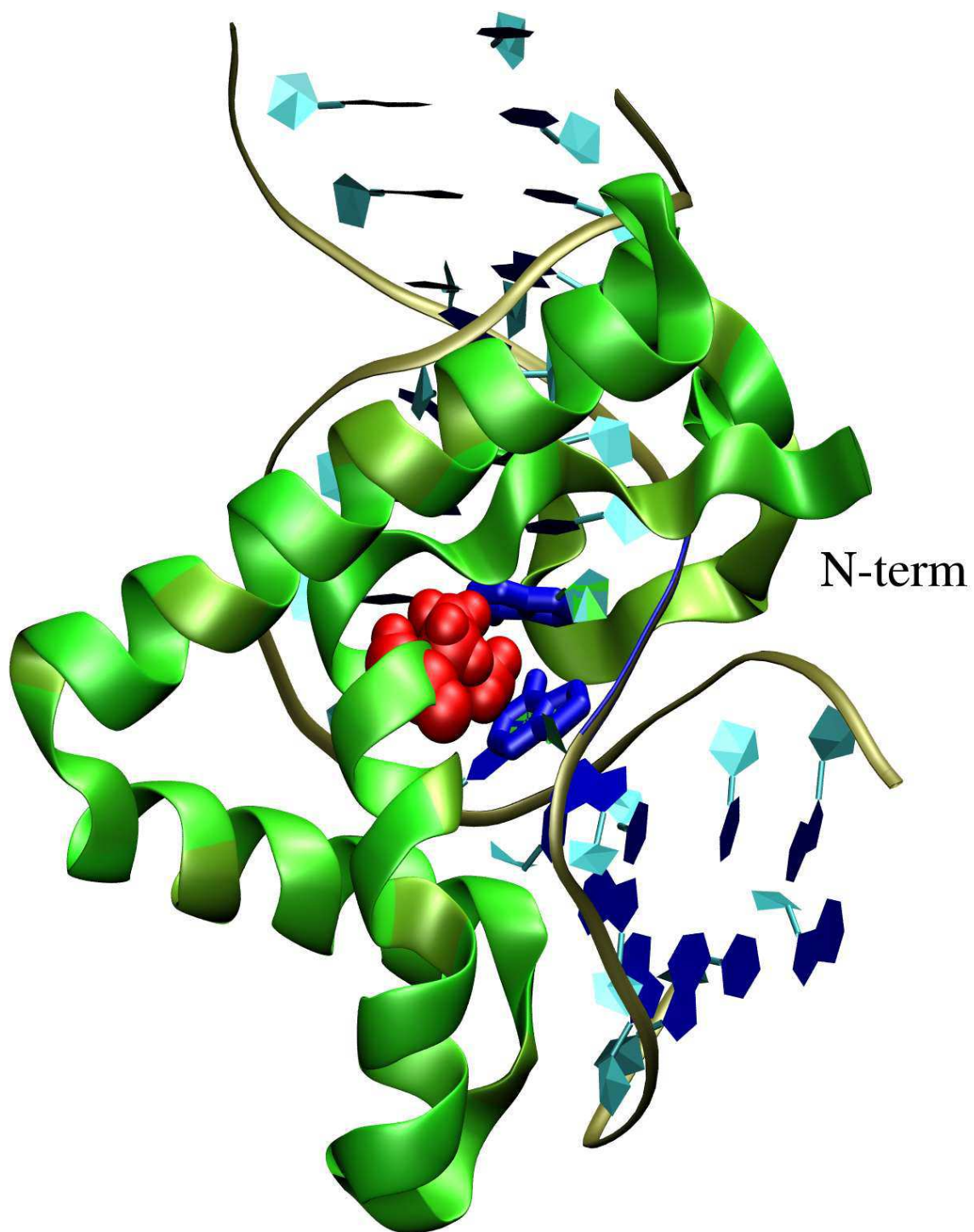


Figure 1

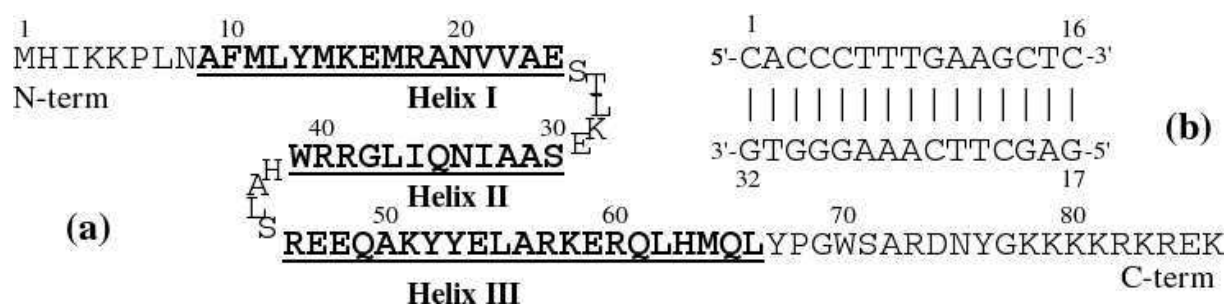


Figure 2

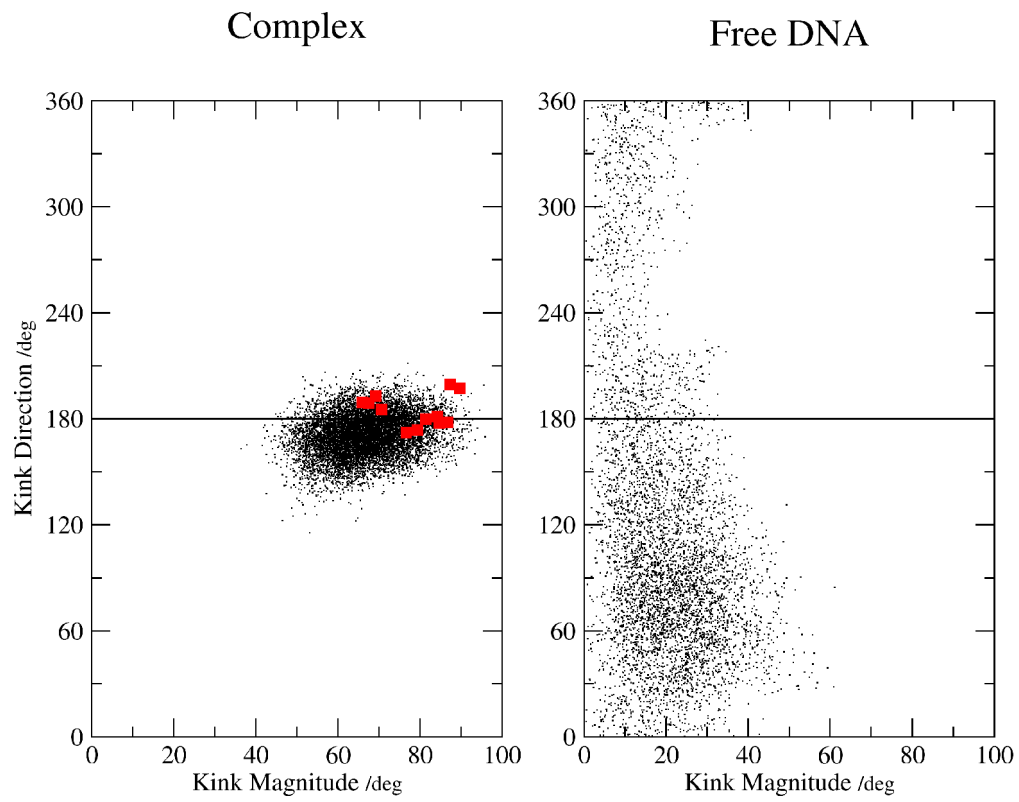


Figure 3

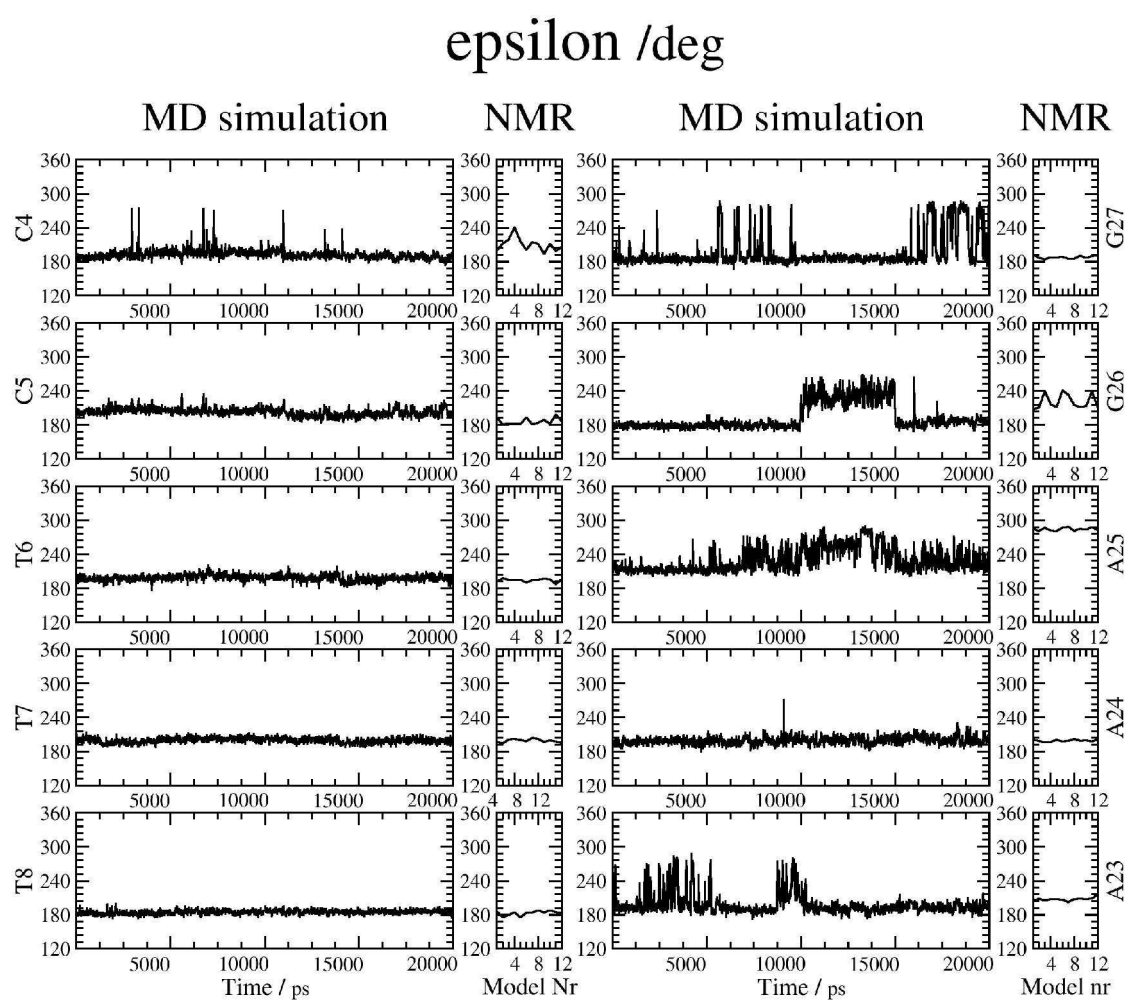


Figure 4a

zeta / deg

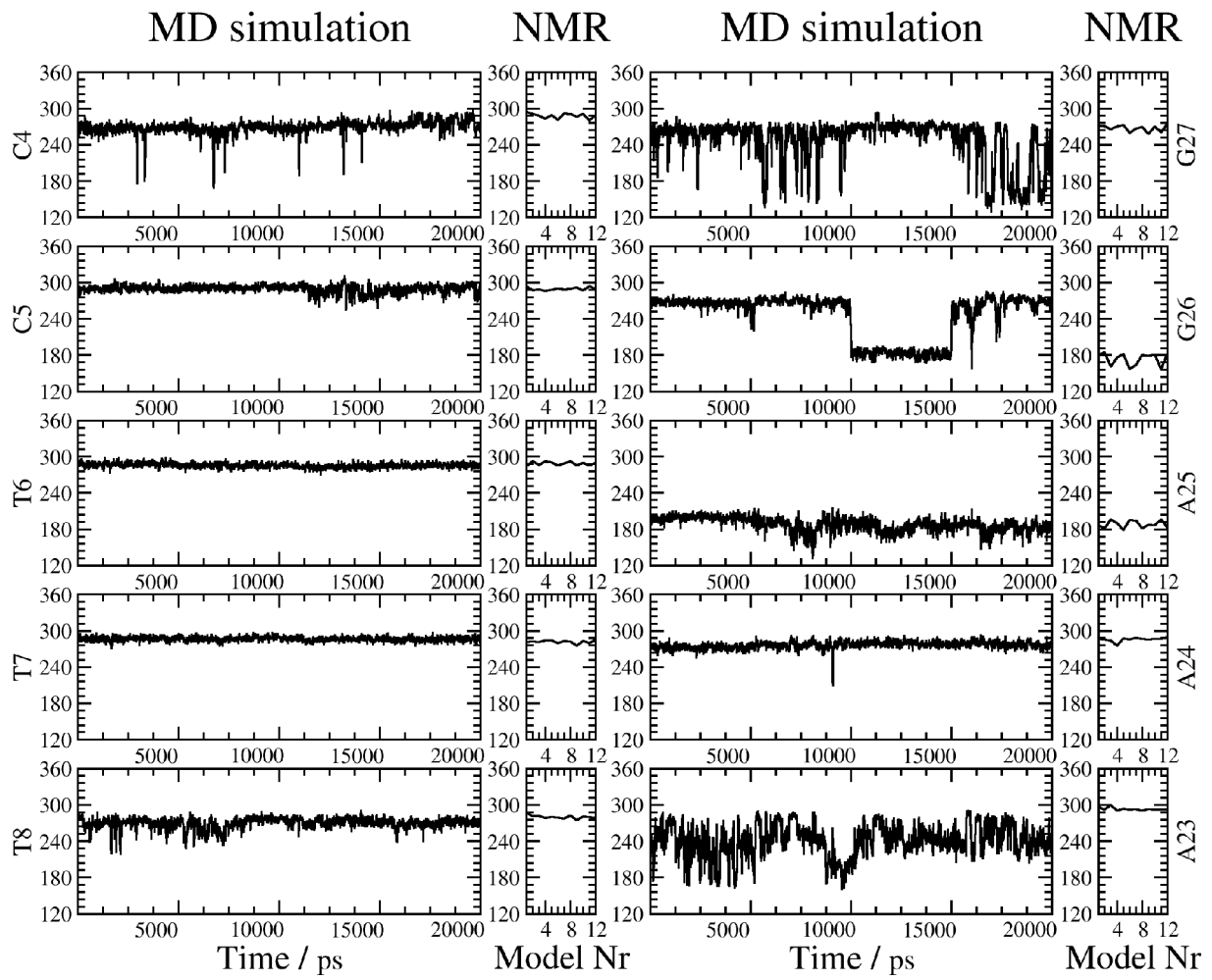


Figure 4b

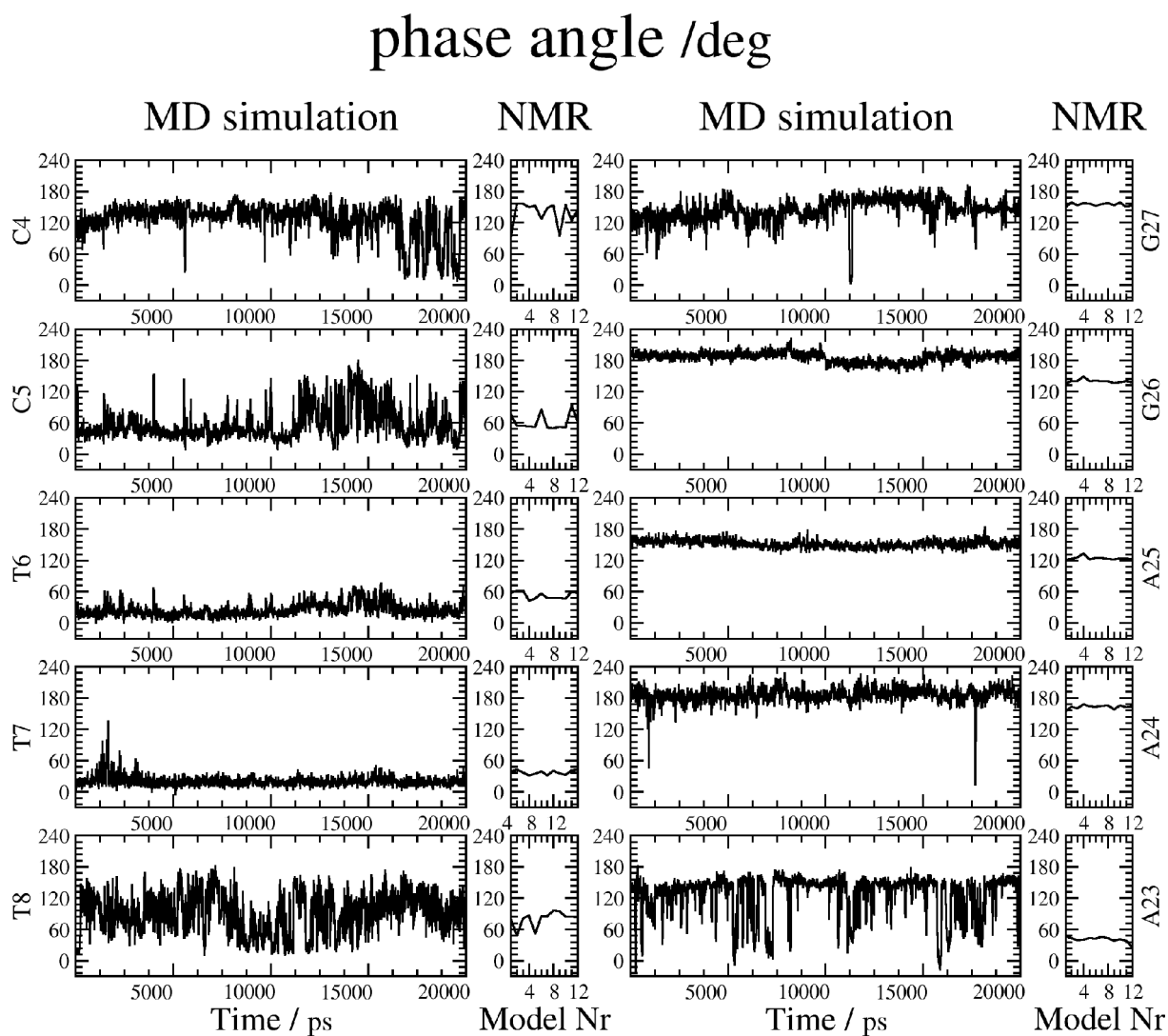


Figure 4c

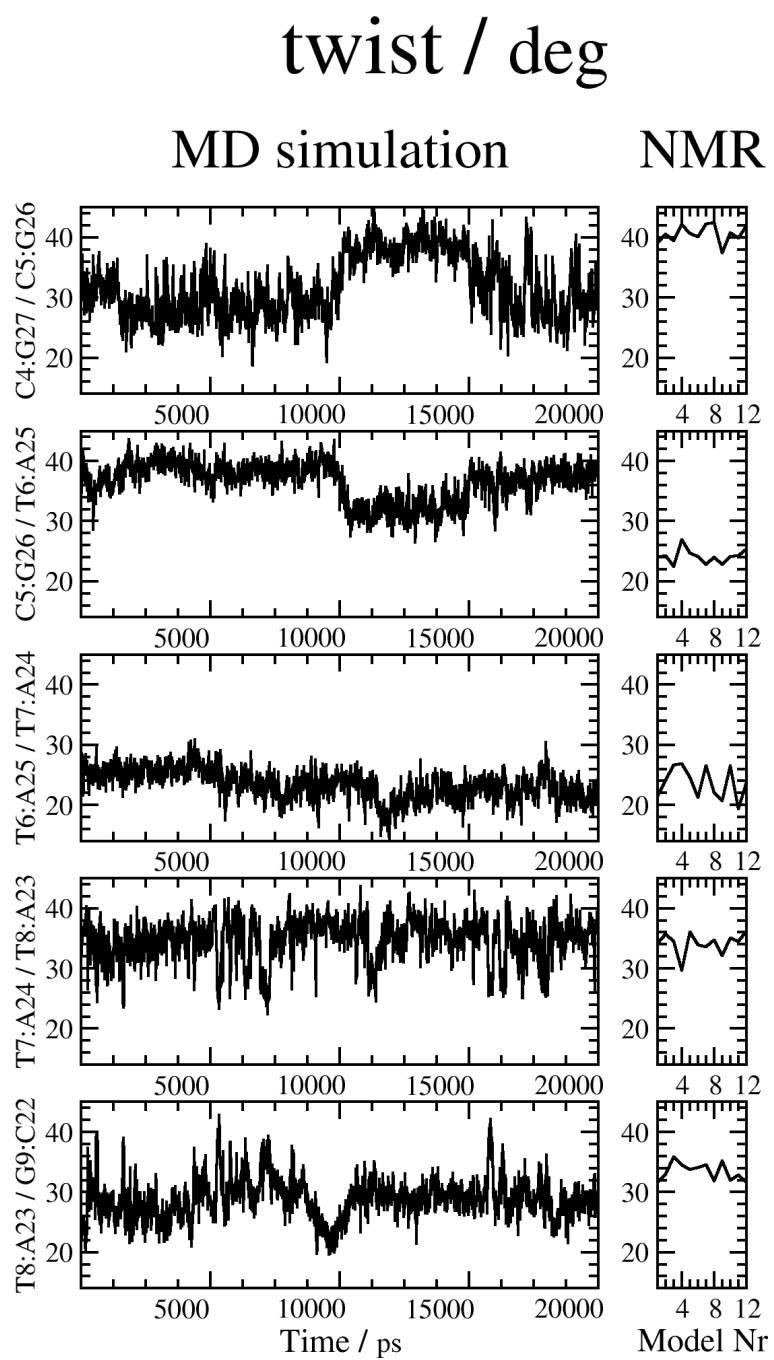


Figure 5a

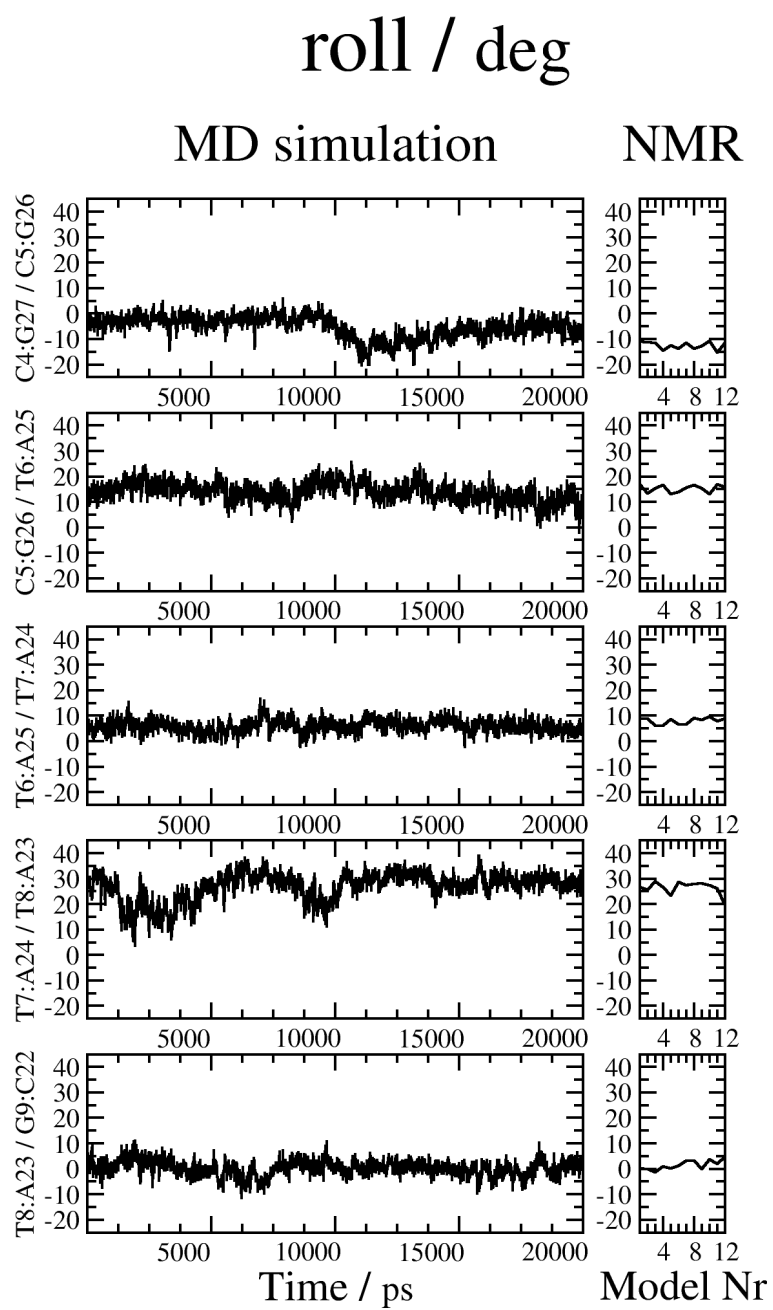


Figure 5b

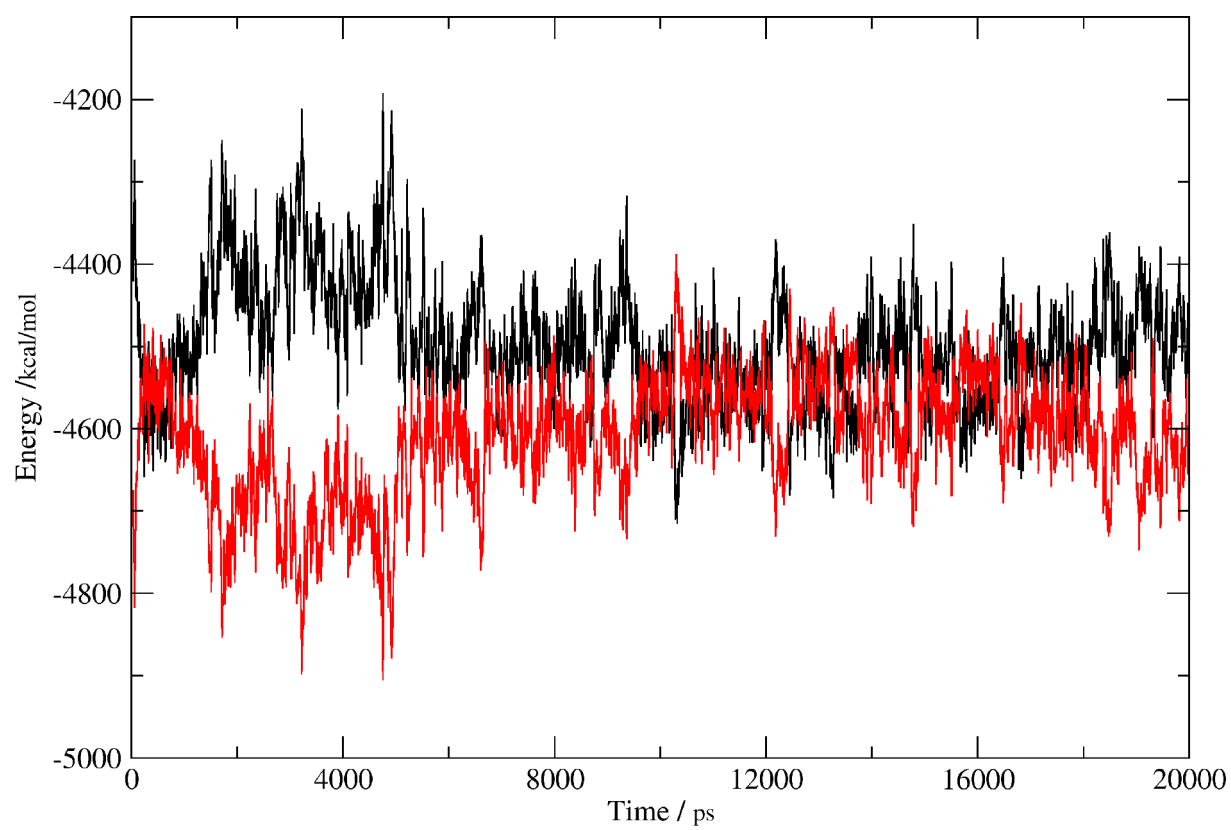


Figure 6

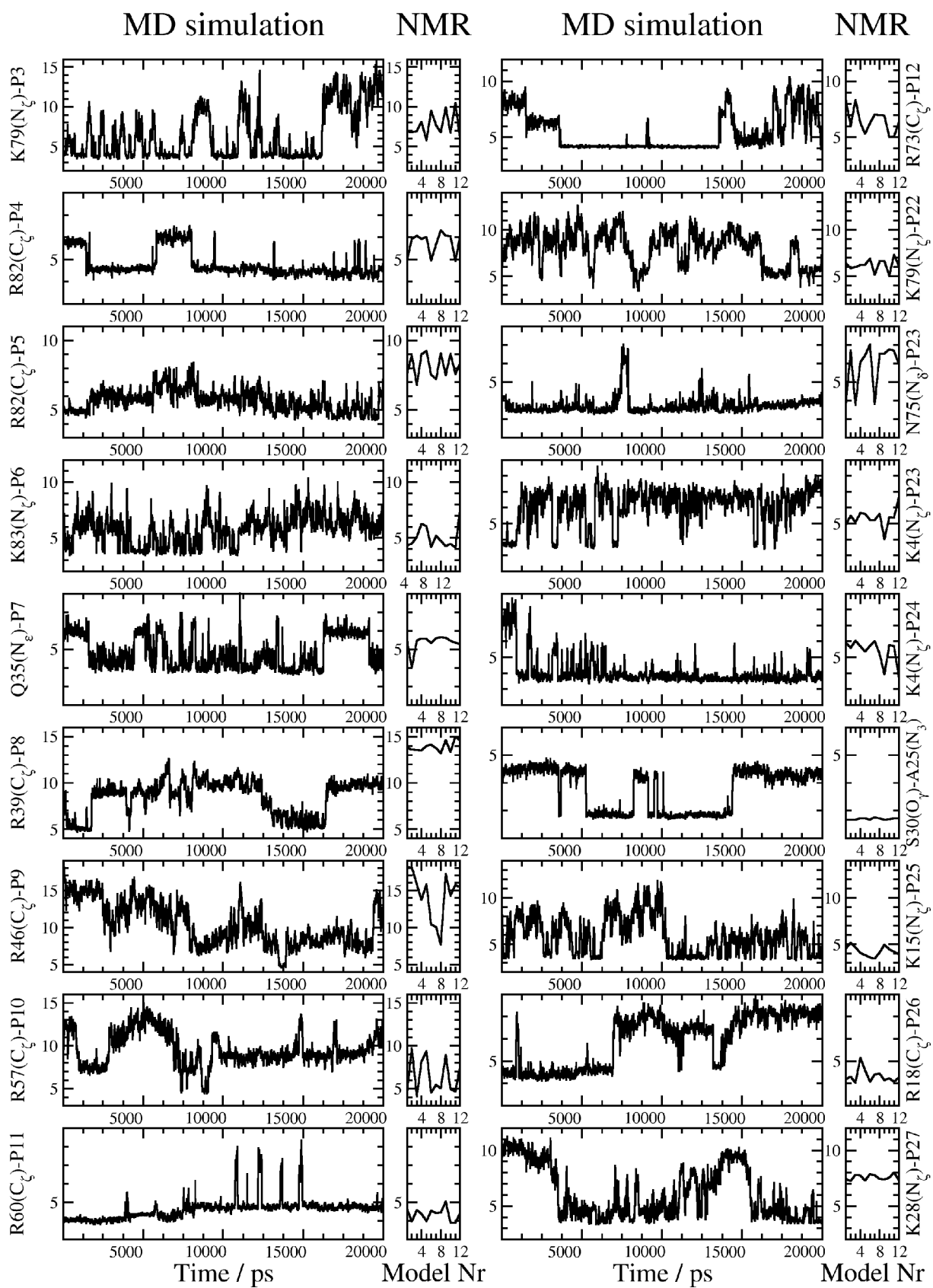


Figure 7a

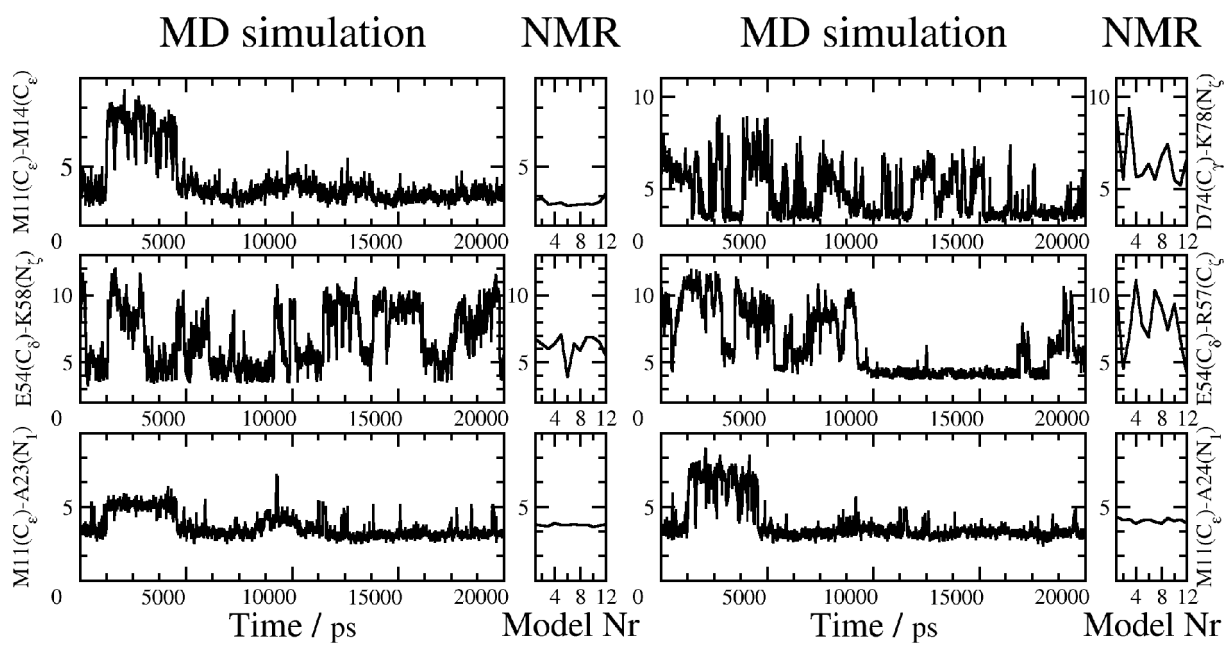


Figure 7b

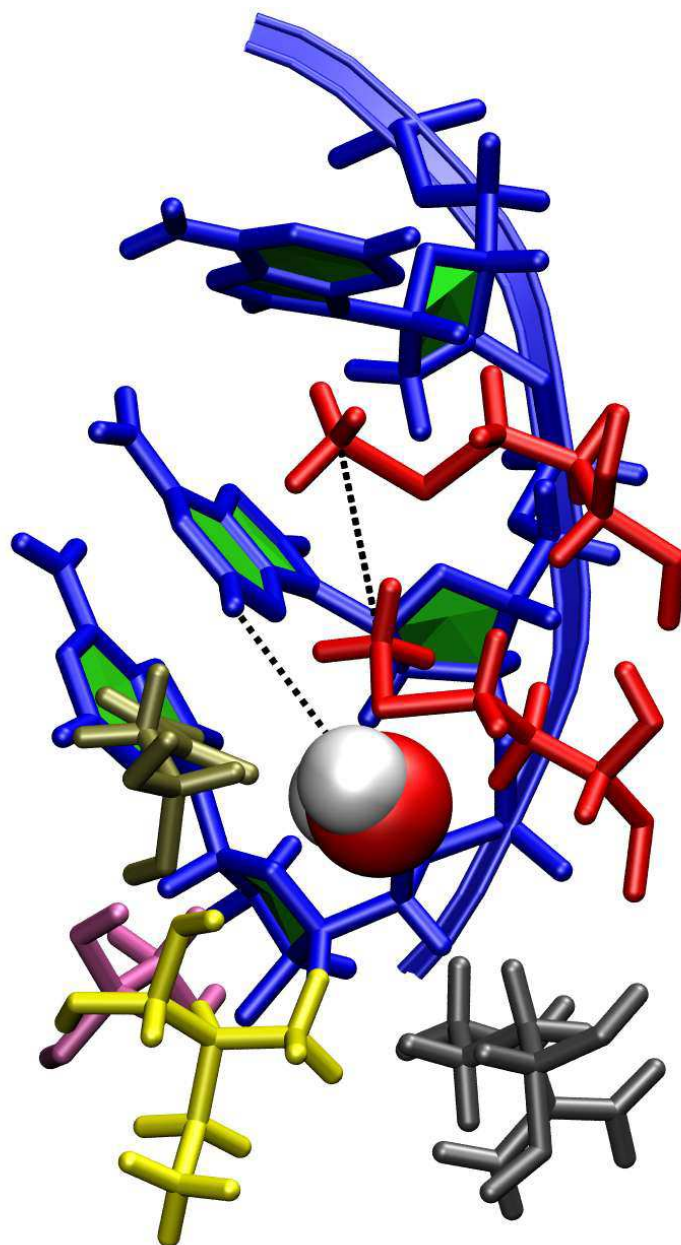


Figure 8a

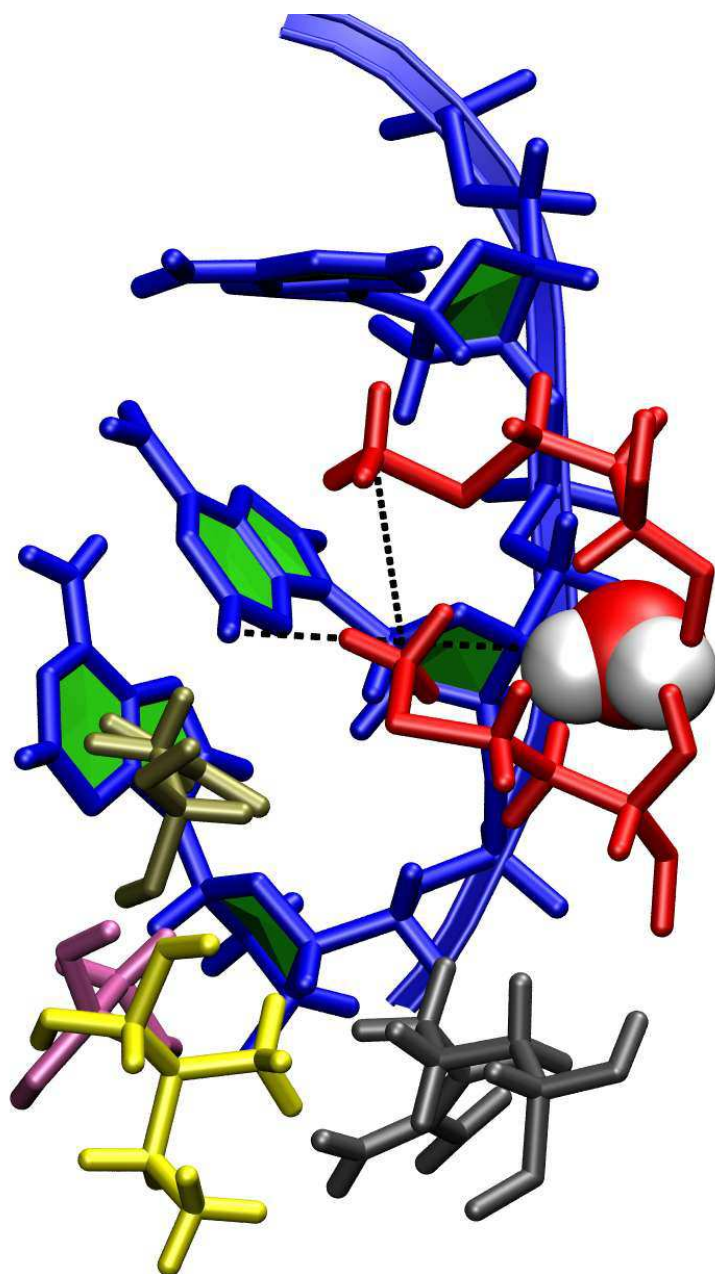


Figure 8b

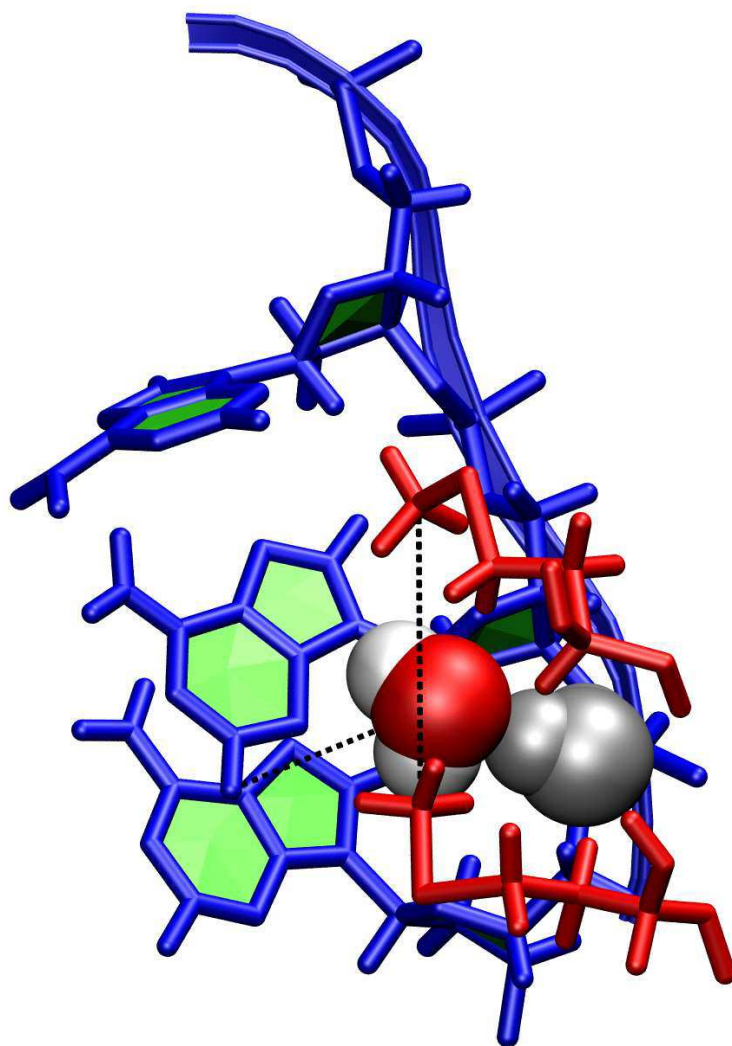


Figure 8c

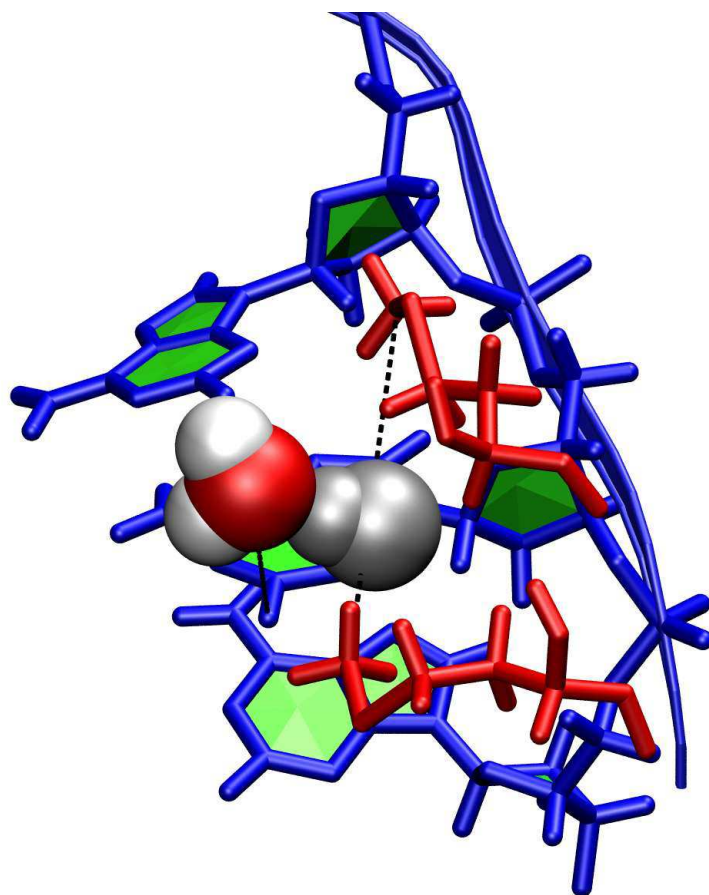


Figure 8d

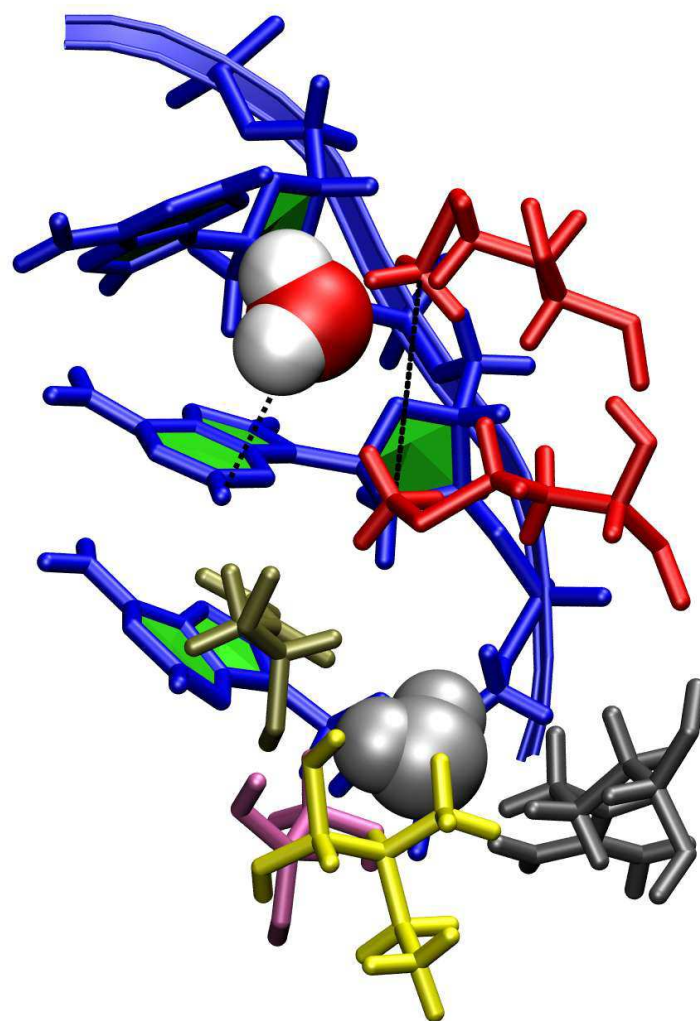


Figure 8e

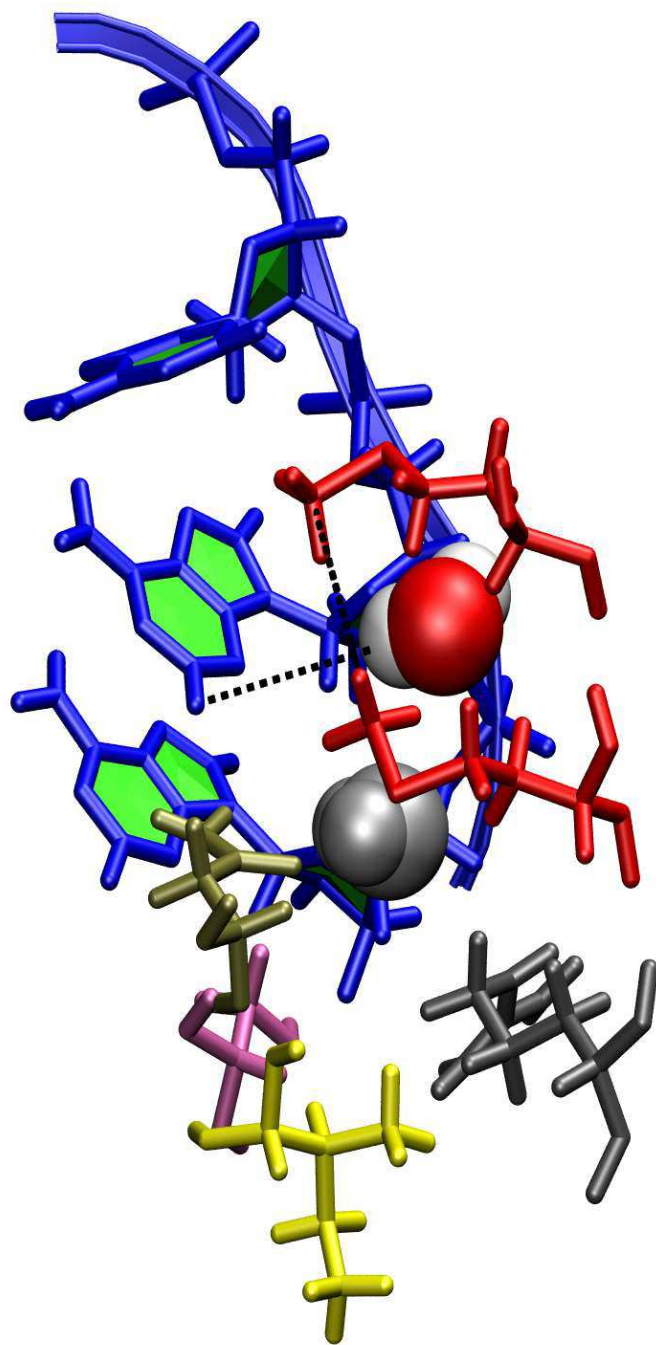


Figure 8f

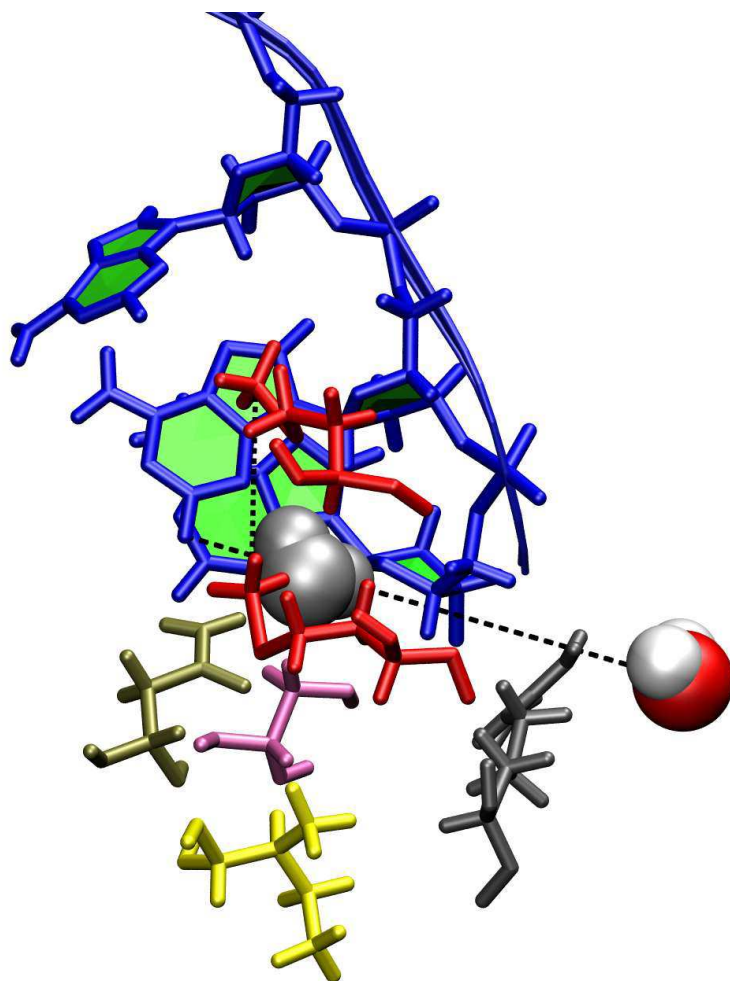


Figure 8g

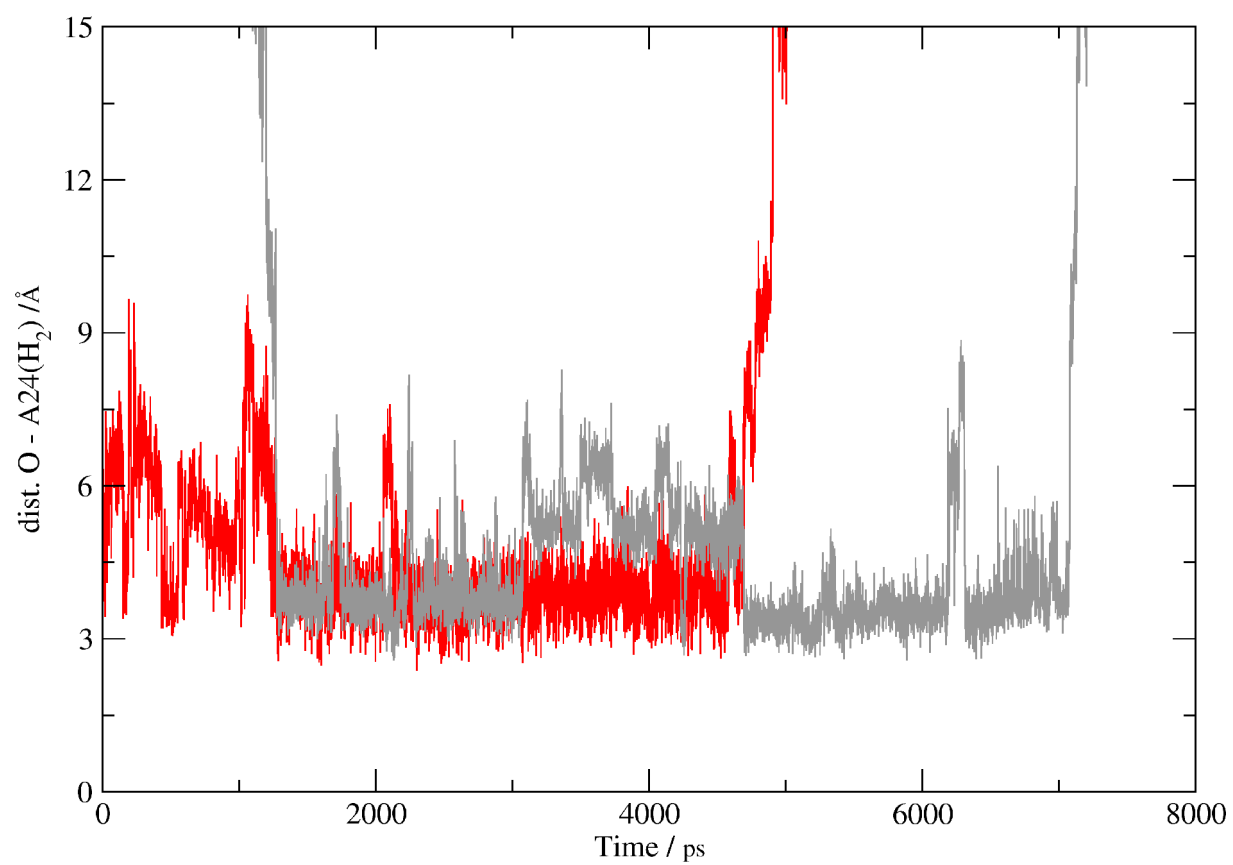


Figure 9

Table SI:

<i>Solute</i>	<i>Number of solute atoms</i>	<i>Number of solvent molecules</i>	<i>Box size / Å</i>	<i>Time / ns</i>
Protein-DNA-Complex	2433 + 16 Na⁺	9427	x: 74,7 y: 73.7 z: 63.7	20.0
DNA 15 nat	950 + 28Na⁺	6571	x: 79,2 y: 59,6 z: 49,5	5.0
DNA 15 gen	950 + 28Na⁺	5410	x: 79,6 y: 49,6 z: 49,5	5.0
DNA 15 GGG	947 + 28 Na⁺	5374	x: 76.2 y: 47.4 z: 47.4	5.0
Protein	1483 + 12 Cl⁻	8384	x: 72.5 y: 73.3 z: 63.4	4.0

Quatrième partie

Etude du complexe
pyrazolato-bisplatine

Les cibles thérapeutiques principales du cisplatine sont restreintes, et bien qu'il soit utilisé en complément de nombreuses chimiothérapies, la recherche d'agents chimiques à spectre d'activité différent est nécessaire. Dans cette perspective les composés multiplatinés ont été étudiés [62, 63, 89, 90]. Ces complexes ont pour caractéristique principale la coordination sur l'ADN avec une distance entre les deux points de platination supérieure à celle du cisplatine, ce qui induit peu de déformations structurales.

Ces complexes ont cependant le défaut de ne pas être directionnels, ce qui conduit à un mélange intra-brin / inter-brins de chélates, à des distances de platination variables.

Pour spécifiquement induire une coordination ciblée à la géométrie fixe, différente de celle du cisplatine, le complexe pyrazolato-bisplatine a été conçu et synthétisé. Celui-ci présente une activité cytotoxique *in vitro*, avec un spectre d'activité différent du cisplatine.

L'équipe du Pr. Reedijk a synthétisé le complexe et procédé à son analyse en RMN. Le travail de simulation a été initié par le Dr. Teuben et le Dr. Elizondo-Riojas (dynamiques préliminaires et inclusion du complexe dans le champ de force) et j'ai effectué les simulations permettant de caractériser le comportement dynamique de l'adduit.

Le modèle issu de la simulation rend compte des observations RMN et valide l'approche théorique de conception d'un composé induisant une déformation différente de celle engendrée par la coordination du cisplatine.

L'étude de la dynamique révèle que le composé déroule son ADN cible d'environ 10° (comme le cisplatine) et que le coude issu de la coordination est similaire à celui de l'ADN libre. La rigidité du complexe ne perturbe pas l'empilement des bases platinées, ce qui explique la non-perturbation de la courbure. En effet la distance $G_5 : N_7 - G_6 : N_7$ de 3,9 Å permet de limiter la déformation base/base.

Cette simulation permet de décrire précisément la structure du complexe pyrazolate-bisplatine. La structure observée présente comme attendu une géométrie régulière de l'ADN (pas de coude prononcé comme dans le cas du cisplatine) et induit localement un déroulement. Ce composé possède une activité cytotoxique *in vitro*, ce qui indique que la déformation induite est suffisante pour induire une cascade réactionnelle. Compte-tenu des connaissances acquises sur le cisplatine, il est raisonnable de supposer que la cascade réactionnelle induite par le composé est différente du cisplatine. Une étude expérimentale permettra de déterminer quelles protéines sont responsables de l'initiation de cette cascade mais il est probable que les protéines HMG n'en feront pas partie.

Le mode d'action du pyrazolate-bisplatine pourrait correspondre à la voie de réparation de l'adduit plutôt qu'à l'induction de l'apoptose.

For submission to *Chemistry- A European Journal*

A pyrazolato-bridged dinuclear platinum(II) complex induces only minor distortions upon DNA-binding

Stéphane Teletchéa,^[a,b] Seiji Komeda,^[a,c,d] Jan-Maarten Teuben,^[c,e] Miguel-Angel Elizondo-Riojas,^[b,f] Jan Reedijk,^{[c]*} and Jiří Kozelka^{[b]*}

^[a] These two authors contributed equally to the present work.

^[b] Dr. J. Kozelka, S. Teletchéa, Dr. M.-A. Elizondo-Riojas
Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques
Université René Descartes, UMR 8601 CNRS
45, rue des Saints-Pères, 75270 Paris Cedex 06 (France)
Fax: (+33)-1-42-86-83-87
E-mail: jiri.kozelka@univ-paris5.fr

^[c] Prof. J. Reedijk, Dr. S. Komeda, Dr. J.-M. Teuben
Leiden Institute of Chemistry, Leiden University
Gorlaeus Laboratories, P.O. Box 9502
2300 RA Leiden (The Netherlands)
Fax: (+31)- 71-527-4671
E-mail: reedijk@leidenuniv.nl

^[d] Present address: School of Chemistry and Biochemistry, Georgia Institute of Technology, 315 Ferst Drive, Atlanta, GA 30332, USA

^[e] Present address: Atlantic Supply, Technical Advisor Fuels, Société des Pétroles Shell, 92708 Colombes Cedex, France

^[f] Present address: Centro Universitario Contra el Cancer (CUCC), Hospital Universitario "Dr. José Eleuterio Gonzalez", Universidad Autónoma de Nuevo León, Monterrey, N.L., Mexico

Abstract: The cytotoxic, pyrazolato-bridged dinuclear platinum(II) complex $[\{cis-Pt(NH_3)_2\}_2(\mu-OH)(\mu-pz)]^{2+}$ (pz = pyrazolate), was found to cross-link two adjacent guanines of a double-stranded DNA decamer without destabilizing the duplex and without changing the directionality of the helix axis. A 1H NMR study of the oligonucleotide d(CTCTG*G*TCTC)-d(GAGACCAGAG), crosslinked at the two G* guanines by $\{cis-Pt(NH_3)_2\}_2(\mu-pz)^{3+}$, and molecular dynamics simulations of the explicitly solvated duplex were accomplished to characterize the structural details of the adduct. The dinuclear platinum crosslink unwinds the helix by approximately 10° , i.e. to a similar extent as the widely used antitumor drug cisplatin, but, in contrast to the latter, induces no significant bend in the helix axis. The Watson-Crick base-pairing remains intact, and the melting temperature of the duplex is unaffected by the cross-link. The helical twist is considerably reduced between the two platinated bases, which becomes manifest in an unusually short sequential H1'-H1' distance. This unwinding also affects the sugar ring of the guanosine 3' to the crosslink which presents an N \leftrightarrow S equilibrium. This is the first antitumor-active platinum complex which was successfully designed envisioning the structural consequences of its binding to DNA.

Keywords: antitumor agents; molecular modeling; nuclear magnetic resonance; oligonucleotides; platinum

Introduction

The structural distortion of DNA caused by the binding of cisplatin (*cis*-diamminedichloroplatinum(II)) is accepted to play a pivotal role in the mechanism of action of this widely used antitumor drug.[1, 2] The clinical use of cisplatin is not only restricted by its toxicity, but also by resistance development, and by the limited scope of cancers against which it is active.[3, 4] Most of the cisplatin analogues that were synthesized to overcome toxicity problems have a similar activity profile as cisplatin.[4-6]

The major cross-link formed when cisplatin reacts with DNA is the 1,2-d(GG)-intrastrand adduct.[7] DNA in this adduct is locally distorted from normal B DNA, and the duplex is bent by 55-78° toward the major groove.[8-14] This distorted Pt-DNA structure is specifically recognized by DNA recognition proteins, which likely mediate the cytotoxic activity of cisplatin.[15-17]

Recent research has focussed on compounds that form different types of cross-links with DNA. Farrell and co-workers[18, 19] have reported various di- and trinuclear platinum complexes that possess antitumor activity *in vitro* and *in vivo* and are characterized by lacking cross-resistance with cisplatin. These compounds produce different kinds of DNA cross-links and the structures of their DNA adducts drugs are drastically different from the distortions induced by cisplatin[20, 21] These results suggest that different DNA binding modes have different biological effects, and thus confer different activity profiles to the drugs inducing them.

One conceivable strategy for the development of new antitumor complexes is based on attempts to diminish the bend angle induced by the cisplatin 1,2-d(GG) cross-link. Following this idea, Kozelka *et al.*[22] have synthesized dinuclear hydrazine-bridged complexes; these species were highly cytotoxic *in vitro*, but failed to show activity in animal models. More promising compounds based on this strategy are azolato-bridged dinuclear platinum complexes recently reported by Komeda *et al.*[23-25] In particular, the complex $[\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-OH})(\mu\text{-pz})]^{2+}$ (pz = pyrazolate, see Figure 1), featuring one pyrazolate and one hydroxide as bridging ligands, was shown to possess a cytotoxicity against MCF7 cells which is ~40 times higher than that of cisplatin, and to be active against cisplatin-resistant cells *in vitro*. [23]

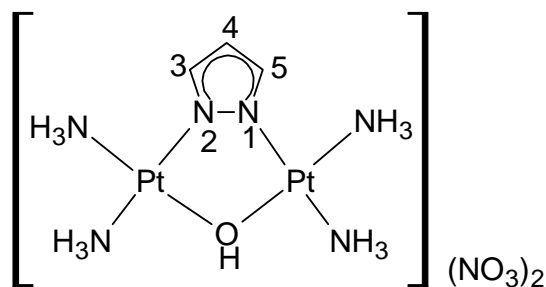


Figure 1. Structure of $[\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-OH})(\mu\text{-pz})][\text{NO}_3]_2$ (**1**). The numbering of the aromatic protons is indicated.

The μ -hydroxide of $[\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-OH})(\mu\text{-pz})][\text{NO}_3]_2$ (**1**) acts as a leaving group, and the bridging rigid pyrazole keeps the distance between the two Pt atoms at the right distance to enable binding of two neighboring guanines. This binding was illustrated in a recently reported crystal structure of the complex $[\{cis\text{-Pt}(\text{NH}_3)_2(9\text{-Et-Gua})\}_2(\mu\text{-pz})]^{3+}$.^[25] In this structure, the two 9-ethylguanines are stacked in a way suggesting that incorporation of such a cross-link into B DNA would cause relatively minor distortions, in contrast to 1,2-GG intrastrand cross-links formed by cisplatin. Thus, **1** is the first antitumor active complex whose design rationale was based on envisioning a specific structural distortion caused upon binding to DNA, i.e., notably avoiding a significant curvature of the DNA helix.

In order to carry out a detailed analysis of this distortion in a DNA duplex, we have reacted $[\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-OH})(\mu\text{-pz})]^{2+}$ with the decamer oligonucleotide d(CTCTGGTCTC) (ssI), and subsequently annealed the platinated single strand with its complementary strand d(GAGACCAGAG) (ssII). The resulting duplex containing the G-Pt-pz-Pt-G cross-link (dsII) was studied using high resolution ^1H NMR and molecular dynamics (MD) simulations. The NOESY spectra of this platinated oligonucleotide differ only locally from the spectra of the unplatinated, undistorted duplex of the same sequence. The modeling work focussed therefore on a precise characterization of the local distortion. To this end, we performed unrestrained MD simulations of the platinated duplex, solvated in a periodic box of ~ 5000 water molecules. The trajectory of the MD simulation was *a posteriori* compared with the NMR data.

Results

Preparation and stability of the platinated DNA duplex: Upon reaction of [$\{cis-Pt(NH_3)_2\}_2(\mu-OH)(\mu-pz)[NO_3]_2$] with ssI at pH 4.0 one major product is seen in the anion exchange FPL chromatogram. The unplatinated ssI (charge -9) elutes at 11.00 minutes, and the newly formed platinated single strand (ssIII) containing the Pt-pz-Pt 1,2-intrastrand cross-link elutes at 9.02 minutes, in agreement with the reduced negative charge of -6 . As in the case of the reaction with 9-ethylguanine,^[25] and 5'-GMP,^[26] no intermediate species are observed. After 3 days the peak corresponding to ssI had almost completely disappeared and ssIII had formed in 90% purity. After titration of the complementary strand ssII (eluting at 10.57 min), the duplex d(CTCTG*G*TCTC)·d(GAGACCAGAG) (dsII; G* denotes a guanine involved in the Pt-pz-Pt cross-link) forms almost instantaneously. The dsII elutes at 12.29 min.

T_m determination: An important measure of distortion is the melting temperature of the platinated DNA compared to the unplatinated DNA. The melting curves of dsI and dsII were determined at identical conditions ($3.4 \cdot 10^{-9}$ M in 1 M NaCl, 50 mM phosphate buffer, pH 7.0). The $T_m = 49 \pm 1$ °C for the unplatinated dsI and $T_m = 48 \pm 1$ °C for the platinated dsII are within the error limits, and therefore it can be concluded that the DNA duplex is not significantly destabilized by binding of $\{cis-Pt(NH_3)_2\}_2(\mu-pz)^{3+}$.

Duplex stability of the NMR sample of dsII (4 mM, 50 mM phosphate buffer pH 7.0, no additional NaCl added) was checked with NMR before conducting 2D experiments. All exchangeable protons remain visible until the duplex is heated to 293 K. At higher temperatures these signals broaden and at 323 K they disappear entirely. All subsequent NMR experiments were conducted at 283 K. At this temperature the sample is still in duplex conformation and overlap of the H3' peaks with the H₂O peak is minimized. Both dsI and dsII were found to be stable for at least 1 week at ambient temperature.

NMR studies

Assignments of chemical shifts: The labeling of atoms and torsion angles in an oligonucleotide is defined in Figure 2. Combined use of TOCSY and NOESY spectroscopy allowed the assignment of the protons of dsI and dsII.^[27-29] The NOESY region (200 ms mixing time) of duplex dsII

visualizing the cross-peaks of the pyrazole moiety is shown in Figure 3a. The sequential assignment pathway of dsII is shown in Figure 3b.

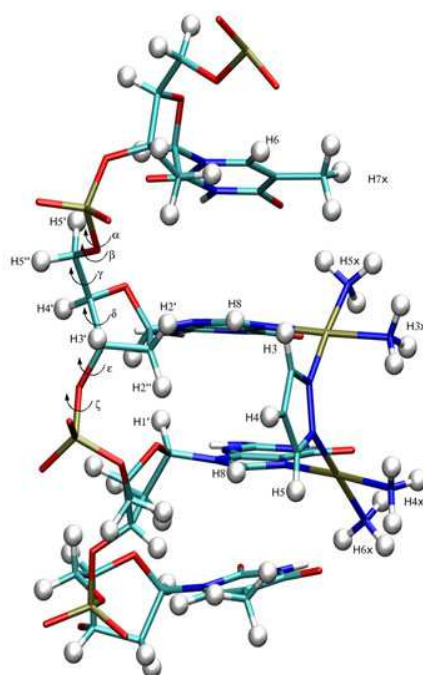


Figure 2. Labeling of atoms and torsion angles within the DNA tetranucleotide TGGT. Crosslinking of the two guanines by the Pt-pz-Pt residue is shown as well. This is a fragment of the structure shown later in Figure 7, and the picture serves to identify atoms in the central part of the platinated strand close in space and susceptible to give rise to NOE crosspeaks.

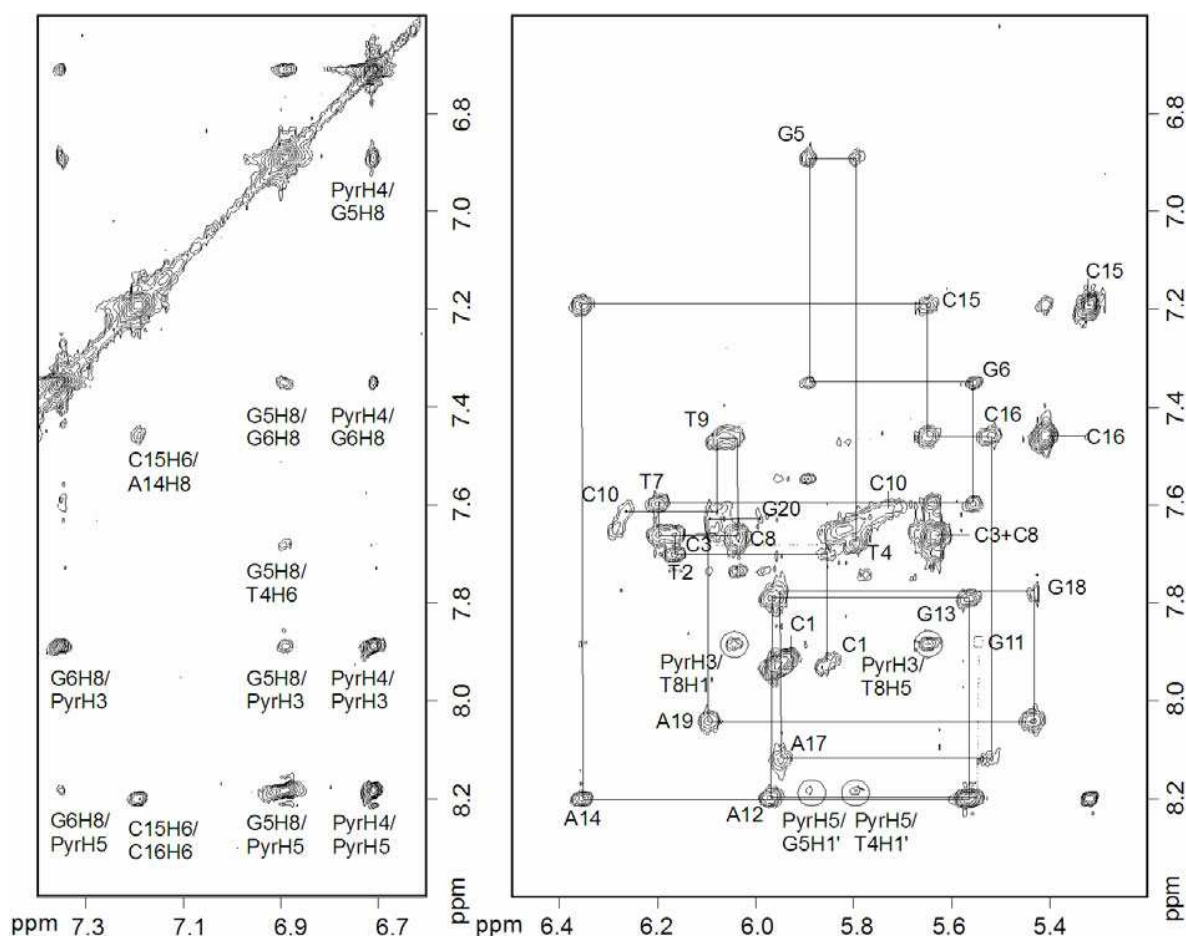


Figure 3. Contour plots of two regions of the 200 ms NOESY spectrum of duplex dsII. **a:** The region of the cross-peaks of the pyrazole protons. **b:** The H8/H6 to H1'/H5 region. The sequential assignment pathway for both strands is indicated. The missing connectivity is indicated with an empty box. Assignments of cytidine H6/H5 cross-peaks are also indicated, and the pyrazole to DNA cross-peaks in this region are encircled.

The adenine H2 protons could easily be identified from their long relaxation times in the T1 relaxation experiment, and were assigned on the basis of intraresidual H2 to H1' cross-peaks and sequential n H2, $(n+1)$ H1' contacts. H2' and H2'' sugar protons were assigned stereospecifically on the basis of the intensities of the H1' to H2'/H2'' contacts.^[30] No stereospecific assignments were obtained for H5' and H5''. For the unplatinated dsI the H2 and H5' and H5'' protons were not assigned. The pyrazole protons (see Figure 1 for numbering) were assigned as follows: the H5 proton was assigned as closest to G5*, *i.e.* the 5'

side of the Pt-pz-Pt lesion. H3 is directed to the 3' side and H4 is the central proton. Figure 3a shows the cross-peaks of the H8/H6 protons with the three pyrazole protons. In the imino region of the ^1H NMR spectrum in H_2O all imino protons were identified (with the exception of those of the 3' terminal G11), indicating that the Watson-Crick hydrogen bonding is conserved among the platinated bases. Each of these imino protons shows cross-peaks in the NOESY spectra in $\text{H}_2\text{O}/\text{D}_2\text{O}$. The chemical shifts of the non exchangeable protons are listed in Table 1.

The plot of differential chemical shifts of the platinated dsII versus those of the unplatinated dsI (Figure 4) immediately revealed some intriguing differences. Presence of the Pt-pz-Pt complex shifts the resonances in the vicinity of the platinum lesion. Most striking is the *upfield* shift of the H8 protons of G5* and G*6. Upon platinum coordination at the N7 atom, the H8 protons are normally shifted *downfield*, because of the inductive effect of platinum binding.^[31] This exceptional upfield shift of G5* and G*6 can be explained by the position of the pyrazole in the major groove (*vide infra*). The protons of the bases T4 and G5* exhibit larger differences in chemical shifts than those of G6* and T7, and the shift differences are minor for the other bases in the duplex. In fact, only minor differences are observed for the sugar protons of the two cytosines opposing the platinated guanines (C15 and C16), but the aromatic protons NH1 and NH2 of these bases are significantly shifted upfield upon platination. The NH1 protons shift 1.08 and 1.53 ppm for C15 and C16 respectively, and the NH2 protons shift 1.41 and 1.03 ppm (see Table 1).

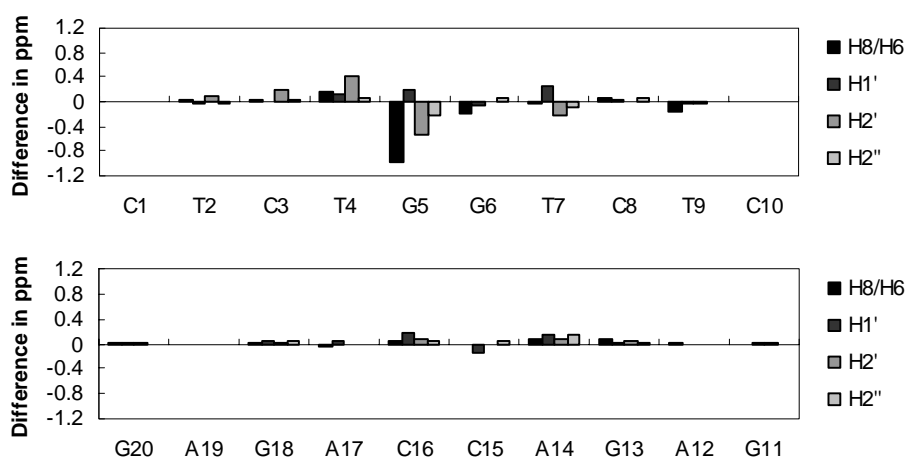


Figure 4. Differences in chemical shift upon platination (dsII-dsI)

NOESY spectra and T1 measurements: A comparison of the NOE intensities of dsII with those of the unplatinated dsI (Figure 5) confirmed the indication from the chemical shifts that the structural perturbation is limited to the close vicinity of the platinum complex. The most obvious difference is the appearance of drug-DNA contacts of the pyrazole ligand with the two central guanines. The intensities of these NOEs differ by factors of up to 30 and indicate that the G*5 base is closer to the pyrazole ring than G*6. The NOE intensities determined at the mixing time of 200 ms correlated well with the time average of the $-1/6$ th power of the internuclear distance over the MD simulation (vide infra). The other NOE cross-peaks were classified as strong, medium, or weak and are summarized, for the central parts of both the platinated (dsII) and unplatinated (dsI) duplexes, in Figure 5. The most striking difference between dsI and dsII is that dsII exhibits an unusual H1' to H1' contact (medium intensity) between G5* and G6*, indicating a decreased distance between these protons.

Table 1. Chemical shifts (ppm) of the protons in dsI and dsII at 283 K

DsI	H8/H6	H2/H5/Me	H1'	H2'	H2''	H3'	H4'	NH1/NH2 ^a	H1/H3
C1	7.94	5.96	5.86	2.29	2.60	4.66	4.11	<i>d</i>	
T2	7.69	1.69	6.21	2.31	2.62	4.66	4.30		13.66
C3	7.65	5.65	6.05	2.13	2.53	4.84	4.20	8.34/7.02	
T4	7.53	1.66	5.67	2.04	2.40	4.85	4.12		13.90
G5	7.87		5.71	2.72	2.75	5.0	4.37		12.97
G6	7.53		5.61	2.21	2.54	<i>c</i>	4.14		12.80
T7	7.64	1.27	5.94	2.52	2.74	4.84	4.40		13.97
C8	7.62	5.61	6.02	2.16	2.53	4.77	4.77	8.40/7.09	
T9	7.63	1.71	6.10	2.18	2.53	<i>c</i>	4.17		14.06
C10	7.66	5.81	6.28	2.27	2.29	4.58	4.05	<i>d</i>	
G11	7.88		5.54	2.50	2.69	4.83	4.16		<i>d</i>
A12	8.19	<i>a</i>	5.97	2.75	2.88	5.06	4.40		
G13	7.72		5.55	2.59	2.71	5.01	4.39		12.65
A14	8.12	<i>a</i>	6.23	2.63	2.91	5.02	4.48		
C15	7.19	5.15	5.78	1.94	2.37	5.15	4.14	8.03/6.73	
C16	7.40	5.15	5.34	1.96	2.28	5.48	4.04	8.48/6.44	
A17	8.17	<i>a</i>	5.90	2.71	2.84	5.03	4.37		12.67
G18	7.74		5.38	2.57	2.66	4.98	4.33		
A19	8.05	<i>a</i>	6.11	2.61	2.90	5.01	4.42		13.36
G20	7.61		5.98	2.36	2.24	4.61	4.15		

Table 1 (continued)

DsII

C1	7.93	5.96	5.86	2.29	2.60	4.65	4.11	7.21/5.97	
T2	7.71	1.67	6.17	2.39	2.59	4.65	4.30		13.94
C3	7.67	5.64	6.04	2.31	2.57	4.78	4.25	8.41/7.08	
T4	7.68	1.73	5.79	2.46	2.47	4.99	4.23		14.08
G5*	6.90		5.89	2.19	2.52	4.99	4.78		13.43
G6*	7.35		5.55	2.20	2.60	5.04	4.27		13.37
T7	7.60	1.18	6.20	2.29	2.65	4.88	4.32		13.70
C8	7.67	5.64	6.04	2.16	2.58	4.84	4.16	8.45/7.18	
T9	7.47	1.73	6.07	2.16	2.53	4.88	4.16		14.04
C10	7.65	5.83	6.28	2.27	2.28	4.59	4.06	8.30/7.24	
G11	7.88		5.57	2.53	2.69	4.83	4.17		<i>d</i>
A12	8.20	7.54	5.97	2.75	2.88	5.06	4.42		
G13	7.79		5.56	2.63	2.73	5.04	4.41		12.76
A14	8.20	7.89	6.36	2.70	3.04	5.04	4.54		
C15	7.19	5.32	5.65	1.92	2.41	4.74	4.11	6.95/5.32	
C16	7.46	5.41	5.52	2.03	2.34	4.83	4.11	6.95/5.41	
A17	8.12	7.79	5.95	2.71	2.83	5.03	4.35		
G18	7.77		5.44	2.59	2.70	5.00	4.35		12.62
A19	8.04	7.74	6.10	2.61	2.90	5.01	4.43		
G20	7.63		5.99	2.38	2.24	4.61	4.16		12.76

^a NH1 is the proton involved in hydrogen bonding. ^c overlap with HDO peak. ^d not observed. G* denotes platination of the guanine.

For G*6, we observe weaker intraresidual H2'-H8 and H2''-H8 NOEs, indicating a shifted conformational equilibrium of the deoxyribose toward a higher population of the N-type conformation. For the S-type conformation normally observed in B DNA, the H8 to H2'/H2'' cross-peaks are strong in intensity, and they are in fact found to be strong for G6 in dsI. Unfortunately, the G6* H3' proton resonates too close to the HDO peak in dsII to allow a reliable quantification of the H8 to H3' cross peak, thus no information on the sugar conformation could be derived from this NOE. However, another indication of an S to N

repuckering of G*6 upon platination comes from spin-lattice relaxation measurements (Table 2) showing that in dsII the T1 value for the H8 proton of G*6 is ~3 times longer than for the other H8 protons. Since the most efficient pathway for H8 relaxation is over the H2' proton,[32, 33] the increase of the T1 of G*6H8 indicates a decrease of the interatomic H8-H2' distance, as observed for an S→N transition. In contrast to G*6, there is no indication of a shifted conformational equilibrium for the other platinated base, G5*.

Another clear difference between dsI and dsII NOESY spectra involves the inter-residual cross-peaks between G6* and T7: the cross-peaks connecting H2' and H2'' of G*6 with the methyl group of T7 have medium intensity in dsII, whereas in dsI, they are weak and almost invisible, respectively. Since within B-DNA, the sequential H2'(n)-TCH₃(n+1) distance is largely independent of the sugar pucker on nucleotide n, the decrease of this distance must be due to an approach between the G*6 deoxyribose and the T7 base, or to a change in the glycosidic torsion angle χ of T7. The distance H2''(n)-TCH₃(n+1), on the other hand, does depend on the sugar pucker of nucleotide n, significantly *increasing* when this nucleotide passes from S to N. The enhancement of the G*6H2''-T7CH₃ NOE upon platination thus argues against a pure N conformation of G*6.

Other differences in NOE intensities upon platination occur in C15 (nucleotide complementary to G6*). The intraresidual H6 to H3' cross-peak increases from medium to strong intensity at the cost of the H6 to H2'/H2'' cross-peaks, which decrease in intensity from strong to medium; the intensity of the H6 to H1' cross peak, however, has a normal intensity. This observation indicates that the mean conformation of the sugar of C15 is also shifted towards a lower phase angle. For comparison, the distance between the H6 and H3' protons is 2.8-3.0 Å for an N-type conformation and 4.0-4.4 Å for the S-type conformation that is normally seen in B DNA.

Table 2. Spin-lattice relaxation times (T1 values) of H8/H6 protons of dsII at 283K

Res.	C1	T2	C3	T4	G5*	G6*	T7	C8	T9	C10
H8/H6 (ppm)	7.93	7.71	7.67	7.68	6.90	7.35	7.60	7.67	7.47	7.65
T1 (s)	1.8	a	1.5	1.0	1.0	3.5	1.0	1.5	1.0	1.0
Res.	G20	A19	G18	A17	C16	C15	A14	G13	A12	G11
H8/H6 ppm	7.63	8.04	7.77	8.12	7.46	7.19	8.20	7.79	8.20	7.88
T1 (s)	0.8	1.1	2.3 a	0.8	1.0	a	a	2.3 a	a	a

^a Overlap of proton signals

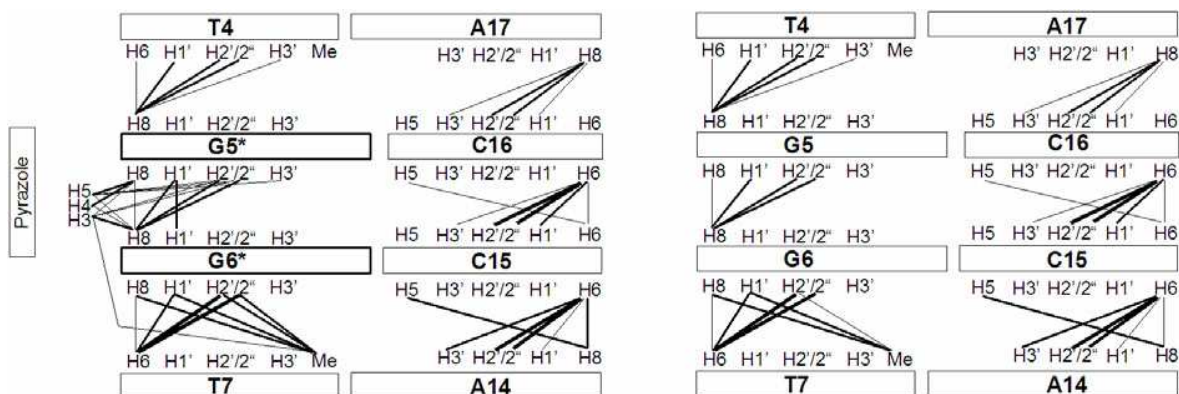


Figure 5. Schematic illustration of the internucleotide NOE cross-peaks in the platinated duplex dsII (left) and the unplatinated duplex dsI (right). The intensities are represented by the thickness of the lines denoting strong, medium and weak cross-peaks.

Molecular Dynamics Simulations of Solvated Duplex dsII

A 20 ns MD simulation of the platinated duplex dsII, explicitly solvated in water, was carried out as outlined in the Experimental Section. Since the above analysis of NMR data suggested that the nucleotides G*6 and C15 might show dynamic S/N equilibria, we were interested in the dynamic features of the structure. We therefore carried out unconstrained MD simulations of dsII, checked that they allowed the salient observations from the NMR spectra (including atypical chemical shifts, NOE intensities, and T1 values) to be rationalized, and used the MD simulation of dsII to predict its global structure. The NMR features of dsI were indicative of normal B-DNA. Our task was thus to find out in what dsII deviates from normal B-DNA double-helix.

The main structural conclusions that one can draw from the comparison of NMR spectra of dsI and dsII (*vide supra*) are:

- i) the structural modifications due to the dinuclear platinum crosslink are limited to the base pairs close to the crosslink;
- ii) the mean sugar phase angles of both G*6 and C15 pass from overwhelmingly S to intermediate between N and S upon platination;
- iii) T7 base seems to be "pushed" closer to the sugar of G*6 by the crosslink;

iv) the pyrazole moiety is somewhat closer to G*5 than to G*6.

All these features are reproduced by the dynamic models, as described in the next paragraphs. The MD trajectories were analyzed using plots of structural parameters (dihedral angles, helicoidal parameters) as a function of time (see Figure S1 for the complete list). The evolution of deoxyribose phase angles and of the helicoidal parameters twist and roll of the four inner base-pairs is shown in Figure 6. The mean values of twist and roll with standard deviations are compared to those of the non-platinated duplex in Table 3.

Different effects of the Pt-pz-Pt crosslink on sugar puckers of G*6 and C15: It can be seen from Figure 6 that the sugar pucker of G*6 undergoes well-defined transitions between N and S type conformations. NMR data (vide supra) have indicated that G*6 presents an N S equilibrium, and the MD simulation is in agreement with this observation. On the other hand, G*5 keeps its S-type pucker, which is also in agreement with the NMR data. This accord indicates that our parameterization of the sugar pucker of the platinated guanines (see Experimental Section) is correct. We recall that these sugar puckers reflect both an inductive effect of platinum binding to guanine (favoring N-pucker^[34-38]) and steric effects. Whereas the former is supposed to affect both guanines equally (and the force-field parameters applied to both platinated guanines were thus identical), the latter are obviously different, and it is this difference in the steric effects that causes the puckers of G*5 and G*6 to be different. The sub-populations of snapshots having either N or S pucker on G*6 were averaged and energy-minimized. The resulting model with G*6N is shown in Figure 7.

The NMR data indicate further that in addition to G*6, the cytidine complementary to G*6, C15, also underwent a change of its sugar conformation upon platination. This nucleotide shows, in the platinated duplex dsII, increased intranucleotide H2'-H6 and H2''-H6 distances, and a decreased H3'-H6 distance. In the framework of the classical two state N/S model,^[39] such changes would be interpreted as indicating a shift of the N S equilibrium towards N. The MD simulation showed instead that C15 assumes an E-type (Est, O4'-endo) pucker. This conformation, characterized by a phase angle of $\sim 100^\circ$, remained stable during the whole simulation, with only a few excursions to S (Figure 6). Thus, the unwinding of the double-helix imposed by the GG crosslink shifts the mean phase angle of both G*6 and C15 towards lower values, but, as the MD simulation indicates, in two different ways: for G*6, by making transitions from S to N more frequent, and for C15, by constraining it to an E conformation.

The reason for this difference could reside in different stabilities of the E conformation in purine versus pyrimidine nucleotides with respect to N and S. In fact, analysis of the high resolution (better than 2 Å) crystal structures extracted from the Nucleic Acids Data Bank, using the selection criteria as described by Djuranovic and Hartmann,^[40] revealed that of the selected 155 AT base pairs, 6% of the adenosines and 28% of the thymidines had an E-type pucker, and of the 233 selected GC pairs, 14% of the guanosines and 22% of the cytidines had an E pucker (E-pucker was defined as $60^\circ \leq P \leq 120^\circ$). Thus, for pyrimidine nucleotides, the E pucker domain seems more readily accessible (i.e., energetically less unfavorable) than for purine nucleotides.

Table 3. Helicoidal parameters Twist and Roll (in degrees) for the central three base-pair steps (calculated with CURVES, using the "global" helix axis), averaged over the MD simulations of the free and platinated duplexes dsI and dsII. Standard deviations are given in parentheses.

Base-pair step	dsI(free)		dsII(platinated)	
	Twist	Roll	Twist	Roll
T4:A17/G*5:C16	29(8)	1(4)	25(8)	4(5)
G*5:C16/G*6/C15	36 (5)	3(6)	23 (5)	3(6)
G*6/C15/T7/A14	30(4)	6(5)	33(4)	7(8)

- Phase -

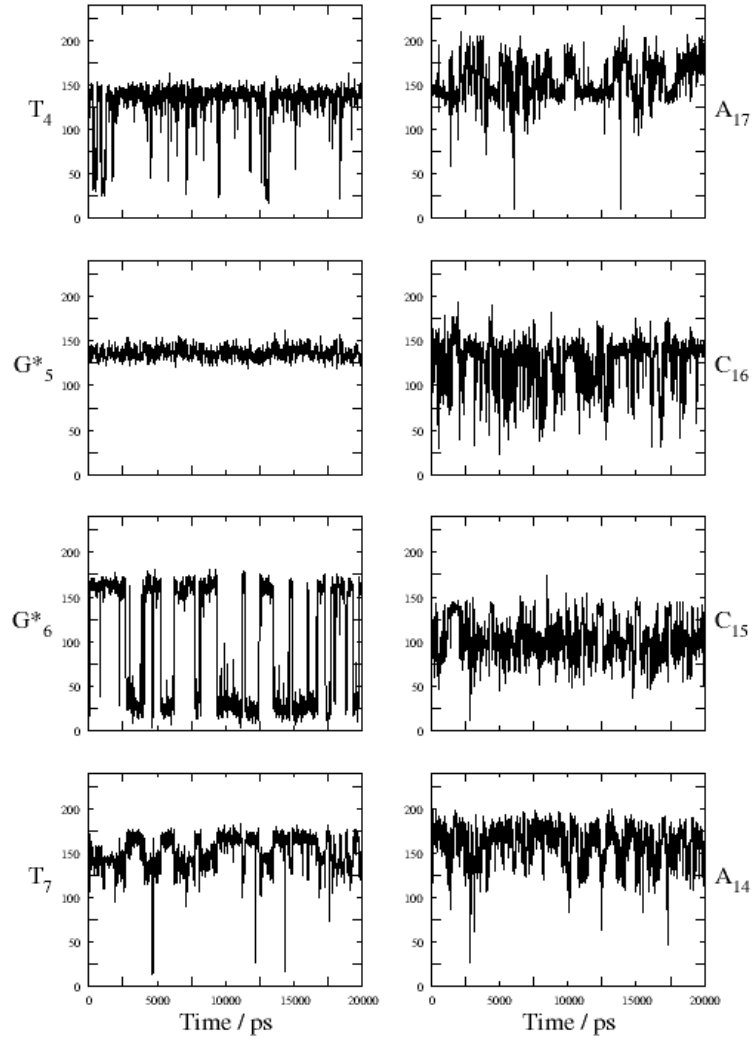


Figure 6a. Deoxyribose phase angles [deg]^[41, 42] calculated for the central four base-pairs of dsII, as a function of simulation time.

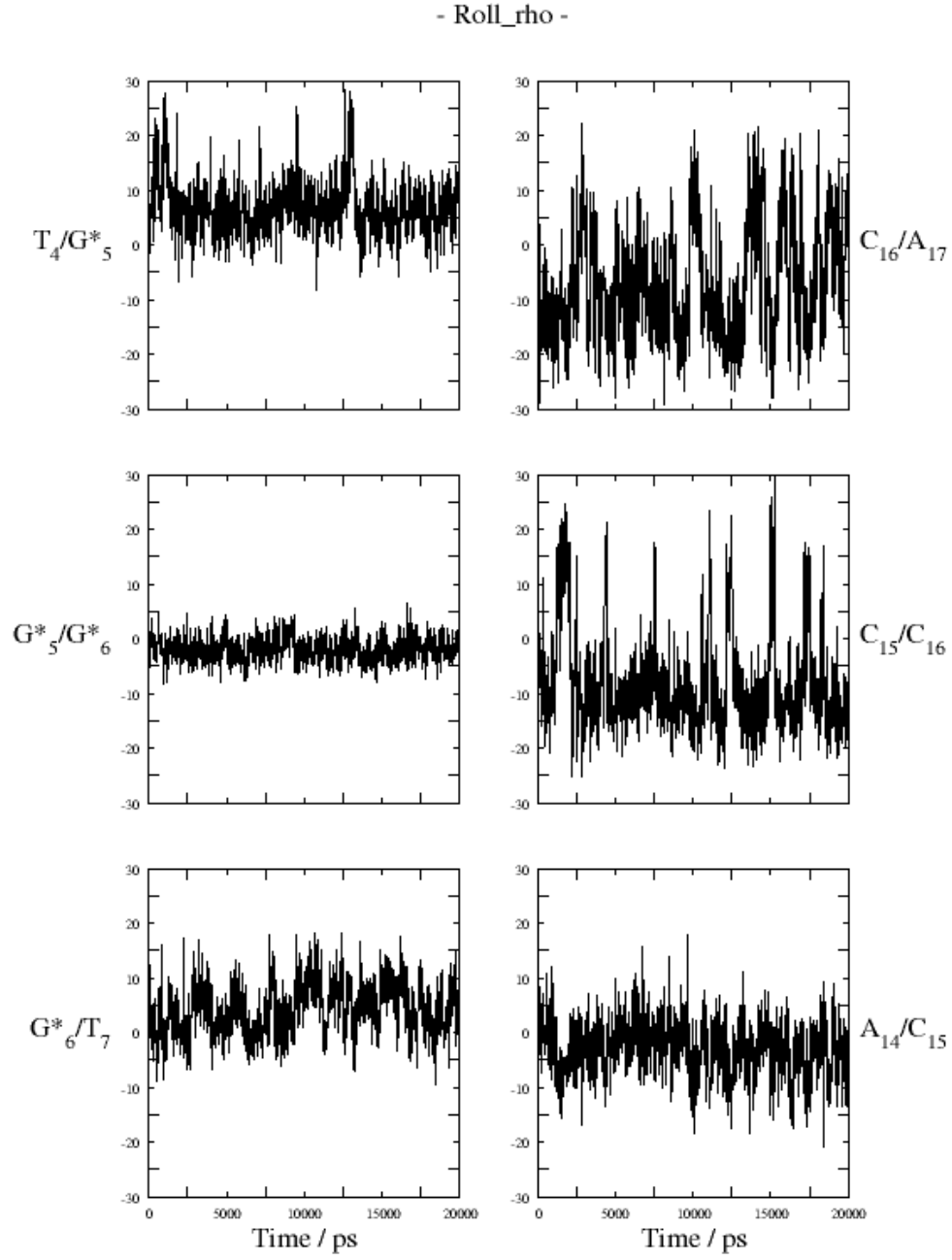


Figure 6b. Base-base roll angles [deg]^[41] calculated for the central four base-pairs of dsII, as a function of simulation time, using the “local” helix axis.^[42]

- Twist_Omega -

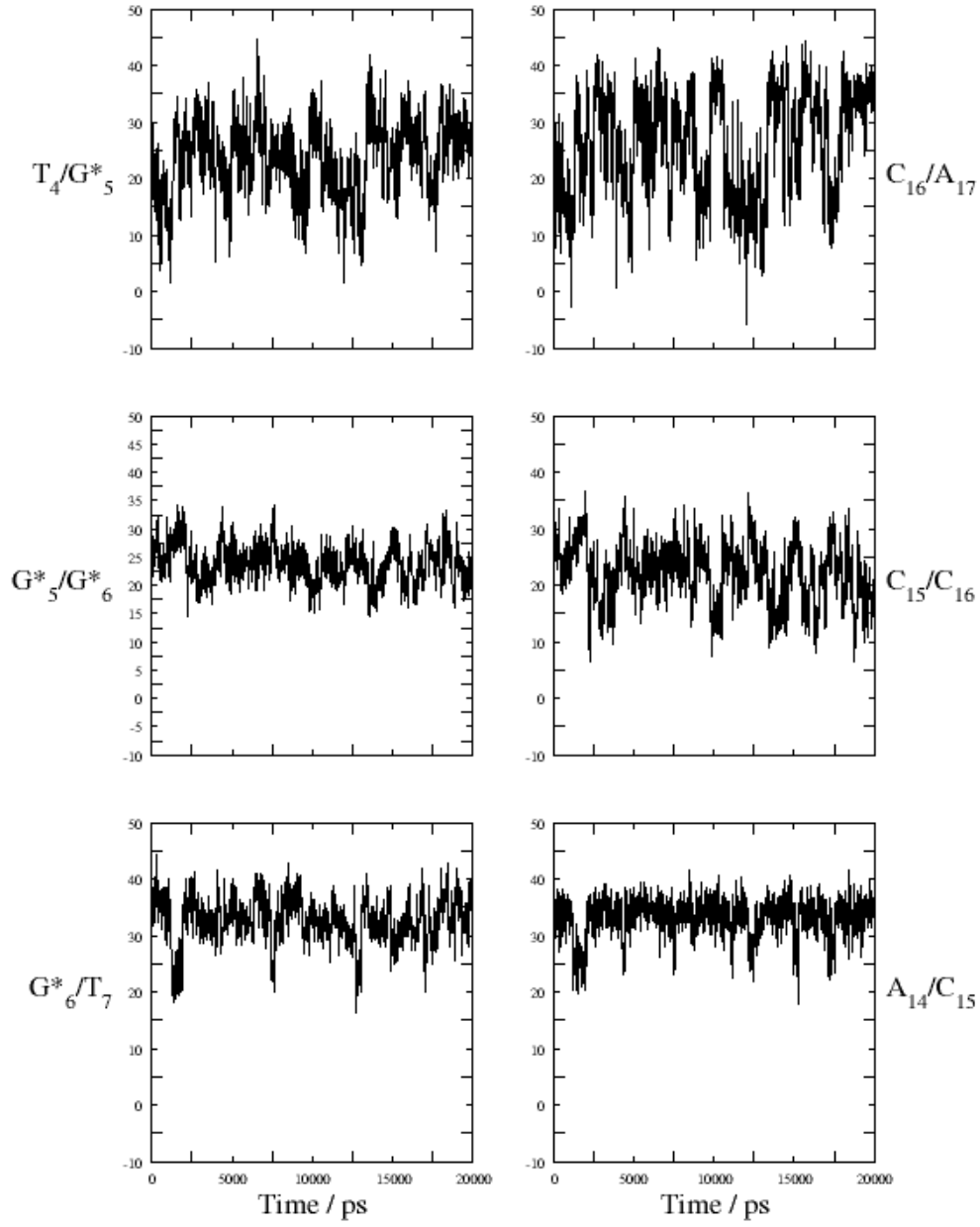


Figure 6c. Base-base twist angles [deg]^[41] calculated for the central four base-pairs of dsII, as a function of simulation time, using the “local” helix axis.^[42]

The Pt-pz-Pt crosslink unwinds the GpG dinucleotide but does not introduce a significant bend: It can be seen from Figure 7 that the platinated duplex dsII does not show significant bending. This is also corroborated by the values of roll (Table 3), which do not change significantly upon platination. On the other hand, the approach of the N7 atoms necessary for the formation of the crosslink causes significant unwinding (diminution of twist) of the crosslinked step G*5:C16/G*6:C15. As seen in Figure 6c and Table 3, this step is unwound from $\sim 36^\circ$ to $\sim 23^\circ$. In the NOESY spectra, this unwinding is reflected by the appearance of the medium intensity G*5H1'-G*6H1' NOE. This distance is observed at 5.1 ± 0.5 Å in the simulation of dsI (not shown), and decreases to 3.2 ± 0.4 Å upon platination. The sugar pucker of G*6 has a significant influence on this distance: the values observed on the time-averaged models are 2.83 and 3.32 Å for N- and S-pucker on G*6, respectively, as can be seen from the detail views presented in Figure 8. This provides a supplementary argument against a pure N-conformation of G*6, since for a distance of 2.8 Å a strong NOE would be expected instead of the observed medium NOE. Table 3 suggests that the unwinding of the G*5:C16/G*6:C15 step is partly compensated by a slight overwinding of the G*6:C15/T7:A14 step. This overwinding is most probably the principal origin of the increase of the intensity of the G*6H2'-T7CH₃ NOE. The time averages of the G*6H2'-T7C7 distance are 3.6 ± 0.4 Å in dsI and 3.3 ± 0.3 Å in dsII.

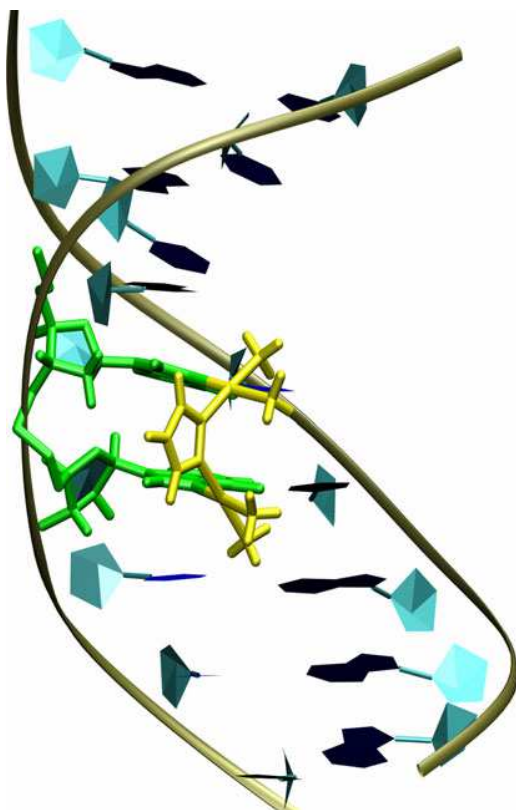


Figure 7. One of the two main conformations observed during the MD simulation of dsII. The model has been obtained by averaging the snapshots having the G*6 sugar in N-type conformation and energy minimization, as described in the Experimental Section. The platinated G*pG* dinucleotide is shown in green sticks and the $\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-pz})^{3+}$ residue in yellow sticks.

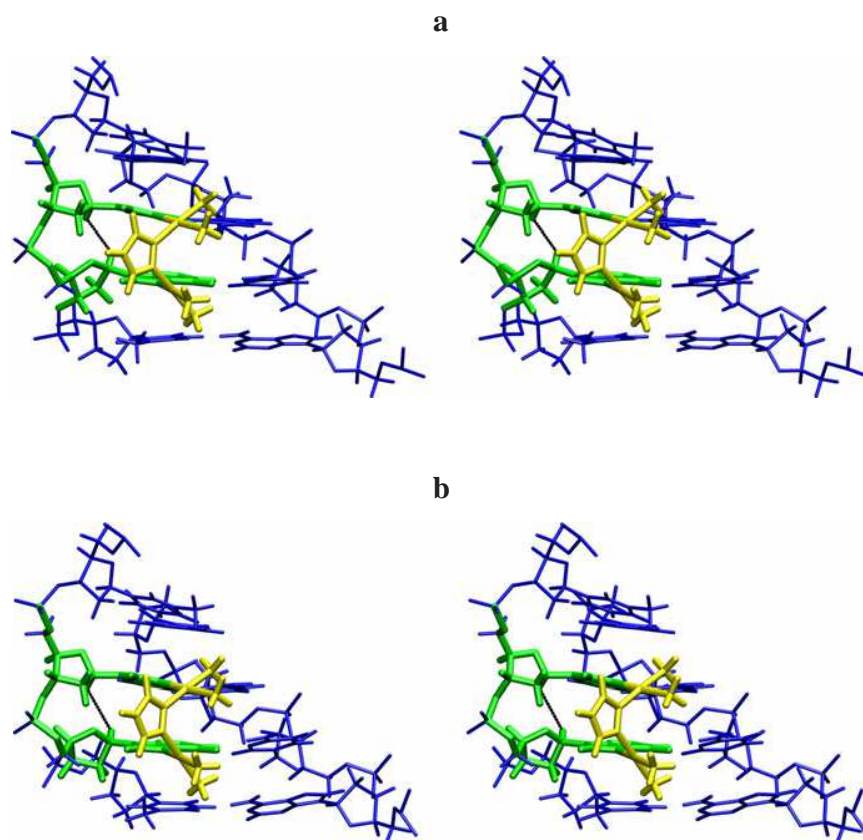


Figure 8. Stereoviews of the inner tetranucleotide d(TG*G*T)-d(ACCA) of the time-averaged models of dsII having G*6 in N (a) and S (b) conformation, respectively. The platinated G*pG* dinucleotide is shown in green sticks and the $\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-pz})^{3+}$ residue in yellow sticks. The G*5H1'-G*6H1' distance is shown as a black dotted line. Its length is 2.83 Å in (a) and 3.32 Å in (b).

Perusal of the backbone torsion angles (Figure S1) shows that the GpG step has assumed the non-canonical $\beta^+\gamma^l$ conformation during the whole simulation. This conformational substate is occasionally found in protein-DNA complexes and is characterized by particularly low twist ($27.4 \pm 0.5^\circ$ as compared to $31.4 \pm 0.5^\circ$ for the canonical $\beta^l\gamma^+$ conformation).^[43] Our simulations indicate that this "strategy" that DNA uses in protein-bound structures to facilitate the unwinding needed for protein recognition is also utilized here to accommodate the unwinding imposed by the covalent crosslink with the dinuclear platinum complex.

Contacts between the platinum ligands and DNA: As can be seen from Figure 8, at both platinum centers the orientations of the guanine with respect to the platinum coordination plane is non-orthogonal; therefore, hydrogen bonds between the NH_3 ligand *trans* to pyrazole and the guanine O6 atom are possible. These N-H...O6 hydrogen bonds were found to persist during the whole simulation, the average N...O separations with standard deviations being 2.90 ± 0.12 and 2.96 ± 0.13 Å for G*5 and G*6, respectively. Moreover, each of these same NH_3 ligands was found to form an additional hydrogen bond to the O4 atom of the neighboring thymine, i.e., T4 with the 5'- NH_3 (3.16 ± 0.26 Å) and T7 with the 3'- NH_3 (3.07 ± 0.17 Å). We have no experimental evidence for these hydrogen bonds; however, their presence is closely related to the position of the pyrazole residue for which we have the experimental check with six NOE distances (vide infra).

The NOESY spectra disclosed six contacts connecting the pyrazole C-H protons and sugar protons of the adjacent residues (G*5, G*6). The NOE intensities determined at 200 ms mixing time are listed in Table 4 together with the corresponding H-H distances time-averaged over the MD simulation. Also given are distances determined from the NOE intensities (I_{NOE}) as $(c/I_{\text{NOE}})^{1/6}$, with c taken as 25000. The excellent agreement between these calculated and NOE-derived values suggests that the platinum binding site is portrayed very well by the simulation. The shorter distances determined for G*5 indicate that the pyrazolate moiety is closer to this base than to G*6.

Table 4. Distances between the pyrazolate and guanine protons, as determined from the NOE intensities, and calculated as time-averages over the MD simulation of dsII.^a

Distance	G*5H8- pzH3	G*5H8- pzH4	G*5H8- pzH5	G*6H8- pzH3	G*6H8- pzH4	G*6H8- pzH5
I_{NOE}^b	1.75	8.38	54.9	20.2	2.75	1.6
$d_{\text{NOE}} [\text{Å}]^c$	4.93	3.79	2.77	3.28	4.57	5.00
$d_{\text{MD}} [\text{Å}]^d$	5.16	4.07	2.30	3.31	4.63	5.04

^a See Figure 2 for numbering of atoms. ^b NOE intensity measured at 200 ms mixing time and given at arbitrary scale. ^c Interproton distance determined from the NOE intensity as $d_{\text{NOE}} = (c/I_{\text{NOE}})^{1/6}$, with c set to 25000. ^d Interproton distance calculated as time-average over the MD simulation of dsII.

Discussion

Chemical shifts and structure: The differences in chemical shifts between the platinated and the unplatinated duplexes contain valuable information about the three-dimensional distortions induced by the platinum complex. These differences are limited to the base-pairs involved in the platinum lesion, and are most pronounced for the aromatic protons of G5*, G6*, C15, and C16. The proximity of the pyrazole ring to the G5*H8 and to the G6*H8 protons (Figure 8) explains—at least in part—the upfield shifts of G5*H8 ($\Delta\delta = -0.97$ ppm) and G6*H8 ($\Delta\delta = -0.18$ ppm) upon platination (see Figure 5).

Platinum coordination to the N7 atom of guanine shifts the H8 proton ~ 0.5 ppm downfield due to the inductive effect of Pt-coordination.^[31] For the H8 protons of G5* and G6* this effect is overcompensated by the shielding effect of the aromatic pyrazole. The fact that the compensation is much more pronounced for G5*H8 is in accord with its position closer to the pyrazole ring.

Another effect probably shifting the H8 resonances of G*5 and G*6 upfield is the unwinding of the duplex, resulting in an increased stacking of the platinated base pairs. This effect of increased stacking is also observed for the amino NH1/NH2 protons of C15 and C16, which are upfield-shifted (by 1-1.5 ppm, Table 1) as well. These protons are positioned too far away from the pyrazole ring to experience any shielding. Weakened Watson-Crick (WC) hydrogen-bonding, which could, in principle, also cause upfield shifts, is not a likely cause here, since the upfield shift is equally observed for the WC and the non-WC protons.

Platinum binding affects the DNA conformation both by sterical constraints and electronic effects: Platinum(II) forms stable and inert covalent bonds to N7 atoms of guanines and adenines. Bifunctional mononuclear Pt(II) complexes, such as cisplatin, are able to crosslink two different purines; such crosslinks cause strong deformations to the DNA double helix. Specifically, cisplatin crosslinks of two adjacent purines bend the helix towards the major groove.^[44] This bending is accompanied by repuckering of some sugar residues, most notably

of that of the 5'-guanosine involved in the crosslink. Such repuckering can be consequence of a local compression of the backbone,[45] but can also have electronic origin, since Pt-bound guanosine nucleotides favor intrinsically the N conformation, as shown already in the 1980's by some of us.[34-37] In fact, coordination of 5'-GMP to **1** has been recently shown to increase the N population in 5'-GMP.[26] In the case of the Pt-pz-Pt crosslink, the geometrical distortion of the double helix is much smaller than in the case of cisplatin crosslinks, nonetheless NMR data indicate a conformational change of the 3'-guanosine towards an N-type pucker. Such a change was not observed in a control MD simulation of dsII with the unmodified parm98 forcefield (the forcefield was only supplemented by parameters defining the energetics of the Pt-pz-Pt moiety, but the torsional parameters of the backbone were unchanged), suggesting that the favoring of N conformation observed for G*6 in the NMR spectra is, at least in part, of electronic origin. We have simulated this electronic effect in the work presented here by decreasing the term $V_3/2$ for the torsion angle OS-CP-CP-OS from 1.5 to 0.75 kcal mol⁻¹ for platinated guanosines. Such a decrease is expected to stabilize the N conformation with respect to S.[46] In fact, this modification shifted the pucker of C*6 from purely S to an N/S equilibrium. Importantly, the G*5 pucker remained S, although the same modification of the forcefield was applied for G*5 and G*6. The structural perturbation (i.e. unwinding) caused by the Pt-pz-Pt crosslink obviously affects G*5 and G*6 unequally, and probably contributes to the S→N repuckering of G*6. The repuckering of G*6 is thus apparently caused by a combination of electronic and steric effects.

The sugar pucker of C15 (base complementary to G*6) is found to be locked in an O4'-endo (E) conformation. Since C15 is on the unplatinated strand, electronic effects of platination are expected to be negligible. The stable E-pucker of C15, a conformation which distinguishes this nucleotide from all other nucleotides of dsII, is therefore quite clearly a consequence of the unwinding of the double helix.

Possible implications for protein recognition: 1,2-GG cisplatin crosslinks are specifically recognized by a number of minor groove-binding proteins, such as the TATA box binding protein,[47] the *E. Coli* DNA repair recognition complex UvrAB,[48] the mismatch repair protein MutS,[49, 50] or HMG box proteins.[15] From a comparison of several DNA-platinum adducts recognized by the chromosomal protein HMG1, Pil and Lippard suggested

that unwinding of the double helix may be an important determinant for HMG1 binding.[51] Helix unwinding widens and flattens the minor groove, and a wide and flat minor groove has indeed been later identified as a recognition element of several HMG box proteins.[16, 52, 53] Since HMG box proteins generally bend the DNA towards the major groove, and pre-bent structures, such as DNA bearing the 1,2-GG cisplatin crosslink, are recognized with high affinity by HMG proteins,[15] the bend could be also considered as a recognition element. However, the question of how the bending and unwinding must be combined to effect HMG binding, is not yet clear. In this respect, the Pt-pz-Pt DNA crosslink, for which dsII is a model, is an interesting test case for HMG recognition, since it causes, as we have shown here, an unwinding of the double helix similar to that of cisplatin, but virtually no bending. In addition to testing whether the Pt-pz-Pt crosslink is recognized by HMG proteins, it would be also of interest to see how the crosslink is recognized by repair proteins and polymerases. Comparing the recognition of the Pt-pz-Pt-DNA adducts with that of cisplatin adducts should provide valuable keys to the recognition mechanism, and hopefully enable the individual contributions of DNA bending and unwinding to protein recognition to be assessed.

Conclusion

The interaction of the new antitumor compound [$\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-OH})(\mu\text{-pz})\}[\text{NO}_3]_2$ with a decamer oligonucleotide was studied using NMR spectroscopy and molecular modeling. The complex was designed with the aim to induce minimal distortions upon binding to adjacent guanines of DNA. We have recently shown that GG sites are favorable targets for this new drug.[54]

The present study demonstrates that the Pt-pz-Pt crosslink of two adjacent guanines induces indeed relatively minor structural perturbations upon the DNA double helix. The DNA is shown to accommodate the crosslink without significant bending or base-pair disruption. Corroborating this result, the melting temperature was found to be identical to that of the unplatinated duplex.

The MD simulation indicates that the only pronounced effect of the crosslink on the global structure is an unwinding by 10-15° of the step between the two platinated GC pairs. This unwinding becomes manifest in an unusual H1'-H1' NOE between G*5 and G*6 and in strong upfield shifts of all amino protons of C15 and C16. The unwinding affects, in turn, the sugars of the G*6:C15 base-pair: that of G*6 shows rapid N-S interconversions, whereas that of C15 is locked in an E-type conformation.

The determination of the structure of the Pt-pz-Pt crosslink at two adjacent guanines was a logical and necessary step towards the elucidation of the antitumor mechanism of the new dinuclear antitumor complex $[\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-OH})(\mu\text{-pz})]^{2+}$. The fact that this complex crosslinks two adjacent guanines of DNA, as cisplatin, induces a similar unwinding, but does not bend the helix axis, makes it a very interesting test compound for future studies of structural determinants governing protein recognition of platinum-DNA adducts.

Finally, we note that $[\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-OH})(\mu\text{-pz})][\text{NO}_3]_2$ is the first successful example of a cytotoxic platinum complex that was designed to produce a specific structural distortion in DNA. It represents a promising and encouraging case of rational drug design.

Experimental Section

Platination Reaction and Sample Preparation: The 10-mer oligonucleotides with the sequence $d(\text{C}_1\text{T}_2\text{C}_3\text{T}_4\text{G}_5\text{G}_6\text{T}_7\text{C}_8\text{T}_9\text{C}_{10})$ (ssI) and its complement $d(\text{G}_{20}\text{A}_{19}\text{G}_{18}\text{A}_{17}\text{C}_{16}\text{C}_{15}\text{A}_{14}\text{G}_{13}\text{A}_{12}\text{G}_{11})$ (ssII) were synthesized by the phosphoramidite method. After purification by preparative anion exchange FPLC (Pharmacia Q Sepharose column, with a linear 0-1.2 M NaCl gradient in 0.02 M NaOH, pH 12) the solution was desalted with a gel filtration column (Pharmacia, Sephadex G25, DNA grade). Purity of the oligonucleotides was assessed using analytical anion exchange FPLC (Pharmacia, Mono Q column); subsequently a Dowex column (Sigma) was used to exchange the counter ion to sodium. FPLC chromatography was performed on a BioLogic HR chromatography system (BioRad).

$[\{cis\text{-Pt}(\text{NH}_3)_2\}_2(\mu\text{-OH})(\mu\text{-pz})][\text{NO}_3]_2$ was synthesized as described^[25] from $[\{cis\text{-Pt}(\text{NH}_3)_2(\mu\text{-OH})\}_2][\text{NO}_3]_2$ and pyrazole. The complex was allowed to react with the purine-rich oligonucleotide (ssI) (6mg) in a 1:1 ratio in 1 ml H_2O . The pH was adjusted to 4.0 with nitric acid. The reaction was monitored over time with FPLC (Pharmacia, Mono Q column, 20 min 0-1.2 M NaCl gradient at pH 12). After 4 days at ambient temperature the reaction was judged to be complete and the reaction mixture was purified using FPLC (Pharmacia Mono Q column, 20 min 0-1.2 M NH_4HCO_3 gradient at pH 9). The platinated DNA solution was subjected to a first round of desalting by rotary evaporation, adding some water, and lyophilization. The residue was redissolved in water and processed to further desalting using a gel filtration column (Pharmacia, Sephadex G25, DNA grade), the desired fractions were collected and lyophilized, yielding 3.5 mg of the platinated single strand $d(\text{C}_1\text{T}_2\text{C}_3\text{T}_4\text{G}_5^*\text{G}_6^*\text{T}_7\text{C}_8\text{T}_9\text{C}_{10})$ (ssIII), containing the 1,2-intrastrand (Pt-Pz-Pt)-DNA cross-link.

The unplatinated duplex dsI was prepared for comparison by titrating ssII to a solution of ssI in D₂O. The annealing was carried out at ambient temperature, and monitored with anion FLPC run at pH 7. The same procedure was repeated with ssII and ssIII to produce the platinated dsII. Both duplexes were lyophilized and redissolved three times with 99.8% D₂O, and finally redissolved in 250 μ l 99.96% D₂O. The solutions were transferred to an NMR tube, dried in a stream of nitrogen and redissolved in 550 μ l of 99.96% D₂O. For H₂O samples, dsII was dissolved in 550 μ l of 90% H₂O/10% D₂O. Final concentrations were 1.0 mM for dsI, and 4 mM for dsII. All samples contained 50 mM phosphate buffer pH 7.0. No additional salt was added.

UV measurements and T_m determination: For concentration determination UV measurements were performed on a Perkin Elmer Lambda 900 UV/VIS/NIR spectrometer. Melting temperatures (T_m) were measured on an Ultrospec 4000 UV spectrometer heating the sample at 1 °C/min from 20 °C to 90 °C, at oligonucleotide concentrations of $3.4 \cdot 10^{-9}$ M in 1 M NaCl and 50 mM phosphate buffer.

NMR spectroscopy: TOCSY and NOESY spectra (using mixing times of 100, 150 and 200 ms) were acquired on a Bruker DMX 600 MHz spectrometer, using the Watergate gradient pulse for minimization of the water signal.^[55] All 2D spectra were collected at 283 K. 512 increments in t₁ were collected, each with 2048 complex data points in t₂, and 64 scans at a sweep width of 6000 Hz. Spectra in H₂O were collected with a sweep width of 12000 Hz. T₁ relaxation experiments were carried out using the standard 180°- τ -90° inversion recovery sequence.

All spectra were processed with FELIX (version 97.0, MSI, San Diego, CA, USA), except for the T₁ relaxation experiments; these were processed using xwinNMR software (Bruker). For 2D spectra the t₁ dimension was zero-filled to 2048 points and a polynomial baseline correction was applied in the t₂ domain of the NOESY spectra. Chemical shifts were referenced to the HDO peak calibrated to a DSS (2,2-dimethyl-2-silapentane sulfonate) standard. The assignment aiding program ANSIG^[56] was employed to aid assignments and to view and compare spectra of the platinated and unplatinated duplex.

Molecular Dynamics Simulations: The MD simulations were carried out with the SANDER module of the program AMBER Version 6.0^[57] run on a cluster of IBM-compatible personal computers under Red Hat GNU/Linux 7.1, or on an IBM Power4 P690 computer operating under AIX 5.1. The Particle-Mesh-Ewald method,^[58, 59] using charge grid spacing of approximately 1 Å with cubic B-spline interpolation and sum tolerance of 10^{-5} Å, was used to calculate the electrostatic energy. A 9 Å cutoff was applied to Lennard-Jones interactions. The MD simulations employed the SHAKE algorithm (tolerance=0.0005) to all X-H bonds^[60] and used a time step of 2 fs.

Additional force-field terms were necessary to parameterize the Pt-pz-Pt moiety. The bond lengths and bond angles were taken from the crystal structure of $[\{cis-Pt(NH_3)_2(9-ethylguanine)\}_2(\mu-pz)]^{3+}$;^[25] and the corresponding force constants were equivalenced to those of the five-membered ring of histidine; for the Pt-N(pyrazole) bonds, the same parameters were used as for the Pt-N(guanine) bonds.^[8] All endocyclic and exocyclic torsion angles not involving Pt were given energy barriers $k_d/2$ in analogy to those used by Amber for the five-membered ring of histidine. The torsion angles X-X-N-Pt which govern the out-of-plane bending of the platinum with respect to the pyrazole were parameterized as follows. First those torsions involving hydrogen atom (H-C-N-Pt) were given a torsion barrier of 1.1 kcal/mol, in analogy to the values used for torsions involving aromatic hydrogen atoms in the Amber database. Secondly the Pt-N-N-Pt torsion was given a barrier of 0, in order to avoid biasing of the mutual positions of the two platinum centers. Thus the principal force keeping the platinum atoms in the pyrazole plane was dependent on the endocyclic torsion angles C-N-N-Pt and C-C-N-Pt. The force constant for these torsions were determined to be $k_d/2 = 40$ kcal/mol with using density functional theory (DFT) calculations, as described below.

Force constants determining the dihedrals between the pyrazole plane and the Pt coordination plane, and the atomic charges of the $[\{cis-Pt(NH_3)_2\}_2(\mu-pz)(pGpG)]^+$ unit were obtained from DFT calculations performed with the program Gaussian94^[61] implemented on a Cray C94 computer. Starting coordinates were taken from the crystal structure of $[\{cis-Pt(NH_3)_2(9-ethylguanine)\}_2(\mu-pz)]^{3+}$, and to save computer time, the 9-ethyl group was replaced by a 9-methyl group. In order to eliminate the influence of crystal packing forces, the geometry was relaxed by minimizing the energy as a function of the four dihedrals along the Pt-N bonds, using Hartree-Fock (HF) calculations. Subsequent DFT calculations employing the hybrid B3LYP functional were

performed on the HF-optimized geometry, using MO coefficients from the HF calculation as trial vectors. The atomic charges were determined by fits to the electrostatic potential obtained from the converged DFT density matrix, using the Merz-Kollman routine implemented in Gaussian94. Both the HF and the B3LYP calculations used the Los Alamos pseudopotential/pseudorbital basis set LANL2DZ.[62] To determine the force constants governing the out of plane bending of the pyrazole, DFT calculation were performed on a the model complex $[\text{Pt}(\text{NH}_3)_3(\text{Hpz})]^{2+}$. The Pt-N3 bond was bent from the Pz plane in steps of 5° . Force constants were obtained by fitting the resulting calculated energies to the amber energy function for dihedral angles.

The sugar-phosphate backbone was modified to take into account the influence of the platination upon DNA fixation on the adjacent sugar and phosphate residues:[34-38] The $V_{3/2}$ potential for the torsion angle O4'-C4'-C3'-O3' was reduced from 1.50 kcal/mol to 0.75 kcal/mol and a $V_{2/2}$ torsion term for the dihedral angle C2'-C3'-O3'-P of 1.00 kcal/mol was introduced to reproduce the stabilization of the ϵ' conformation of platinated guanosines.

Simulation protocol: Using the program NUCGEN, the starting structure for dsII, $d(\text{C}_1\text{T}_2\text{C}_3\text{T}_4\text{G}_5^*\text{G}_6^*\text{T}_7\text{C}_8\text{T}_9\text{C}_{10}) \cdot d(\text{G}_{20}\text{A}_{19}\text{G}_{18}\text{A}_{17}\text{C}_{16}\text{C}_{15}\text{A}_{14}\text{G}_{13}\text{A}_{12}\text{G}_{11})$ (where G_5^*G_6^* denotes the Pt-pz-Pt cross-link) was generated using a canonical Arnott B-DNA[63] as a starting model. The $\{\text{cis-Pt}(\text{NH}_3)_2\}_2(\mu\text{-pz})^{3+}$ residue was then manually positioned using XLEAP. The resulting DNA-chelate was minimized by 100 steps of conjugate gradient minimization. The system was then heated and equilibrated as described for the simulation protocol of Elizondo-Riojas and Kozelka.[13]

Unrestrained dynamics production was carried out for 20000 ps. The 20000 structures (one per ps) of the MD trajectory were analyzed using the programs Carnal (in the Amber package), CURVES,[42] VMD,[64] XmGrace,(<http://plasma-gate.weizmann.ac.il/Grace/>) and in-house software (<http://www.steetch.org>).

Acknowledgment

We are indebted to Dr. B. Hartmann for constructive discussions and numerous analyses of the NDB database. J.K. and S.T. thank the Association for International Cancer Research (AICR) (grant N° 00-321) and l'Association pour la Recherche contre le Cancer (A.R.C) (grant P01/3/4297) for financial support. S.K is indebted to The Kidani Memorial Trust for a

fellowship. Computer time from the IDRIS computer center of the CNRS and support from COST (Projects D20/001/00 and D20/003/00), enabling scientific exchange with other research groups, are gratefully acknowledged. M.-A. E.-R. received the 1997 Gemini award from the International Precious Metals Institute (IPMI). This research has been financially supported in part by the Council for Chemical Sciences of the Netherlands Organisation for Scientific Research (CW-NWO). The authors are grateful to Johnson & Matthey (Reading, UK) for their generous loan of K_2PtCl_4 .

Supporting Material. Figure S1 showing the evolution of all backbone torsion angles and helicoidal parameters of dsII (excluding the terminal base pairs) during the MD simulation.

References

- [1] Z. Guo, P. J. Sadler, *Angew. Chem. Int. Ed. Engl.* 38 (1999) 1512-1531.
- [2] P. J. O. Dwyer, J. P. Stevenson, S. W. Johnson, in B. Lippert (Ed.): *Cisplatin: Chemistry and Biochemistry of a Leading Anticancer Drug*, Verlag Helvetica Chimica Acta, Zürich 1999, p. 29-69.
- [3] G. Chu, *J. Biol. Chem.* 269 (1994) 787-790.
- [4] M. C. Christian, *Seminars in Oncology* 21 (1992) 720.
- [5] J. Reedijk, *Chem. Commun.* (1996) 801-806.
- [6] J. Reedijk, *Proc. Natl. Acad. Sci. USA* 100 (2003) 3611-3616.
- [7] A. M. J. Fichtinger-Schepman, J. L. van der Veer, J. H. J. Den Hartog, P. H. M. Lohman, J. Reedijk, *Biochemistry* 24 (1985) 707-713.
- [8] F. Herman, J. Kozelka, V. Stoven, E. Guittet, J.-P. Girault, T. Huynh-Dinh, J. Igolen, J.-Y. Lallemand, J.-C. Chottard, *Eur. J. Biochem.* 194 (1990) 119-133.
- [9] D. Yang, S. S. G. E. van Boom, J. Reedijk, J. H. van Boom, A. H.-J. Wang, *Biochemistry* 34 (1995) 12912-12920.
- [10] P. M. Takahara, A. C. Rosenzweig, C. A. Frederick, S. J. Lippard, *Nature* 377 (1995) 649-652.
- [11] A. Gelasco, S. J. Lippard, *Biochemistry* 37 (1998) 9230-9239.
- [12] J. A. Parkinson, Y. Chen, Z. Guo, S. J. Berners-Price, T. Brown, P. J. Sadler, *Chem. Eur. J.* 6 (2000) 3636-3644.

- [13] M.-A. Elizondo-Riojas, J. Kozelka, *J. Mol. Biol.* **314** (2001) 1227-1243.
- [14] L. G. Marzilli, J. S. Saad, Z. Kuklenyik, K. A. Keating, Y. Xu, *J. Am. Chem. Soc.* **123** (2001) 2764-2770.
- [15] D. B. Zamble, S. J. Lippard, in B. Lippert (Ed.): *Cisplatin: Chemistry and Biochemistry of a Leading Anticancer Drug*, Verlag Helvetica Chimica Acta, Zürich 1999, p. 73-110.
- [16] U. M. Ohndorf, M. A. Rould, Q. He, C. O. Pabo, S. J. Lippard, *Nature* **399** (1999) 708-712.
- [17] M. A. Fuertes, J. Castilla, C. Alonso, J. M. Pérez, *Curr. Med. Chem.* **10** (2003) 257-266.
- [18] N. Farrell, Y. Qu, U. Bierbach, M. Valsecchi, E. Menta, in B. Lippert (Ed.): *Cisplatin: Chemistry and Biochemistry of a Leading Anticancer Drug*, Verlag Helvetica Chimica Acta, Zürich 1999, p. 479-521.
- [19] Y. Zou, B. Van Houten, N. Farrell, *Biochemistry* **33** (1994) 5404-5410.
- [20] Y. Qu, M. J. Bloemink, J. Reedijk, T. W. Hambley, N. Farrell, *J. Am. Chem. Soc.* **118** (1996) 9307-9313.
- [21] D. Yang, S. S. G. E. van Boom, J. Reedijk, J. H. van Boom, N. Farrell, A. H.-J. Wang, *Nature Struct. Biol.* **2** (1995) 577-586.
- [22] J. Kozelka, E. Segal, C. Bois, *J. Inorg. Biochem.* **47** (1992) 67-80.
- [23] S. Komeda, M. Lutz, A. L. Spek, M. Chikuma, J. Reedijk, *Inorg. Chem.* **39** (2000) 4230-4236.
- [24] S. Komeda, M. Lutz, M. Spek, A. L. , Y. Yamanaka, T. Sato, M. Chikuma, J. Reedijk, *J. Am. Chem. Soc.* **124** (2002) 4738-4746.
- [25] S. Komeda, H. Ohishi, H. Yamane, M. Harikawa, K. Sakaguchi, M. Chikuma, *J. Chem. Soc. Dalton Trans.* (1999) 2959-2962.
- [26] S. Komeda, H. Yamane, M. Chikuma, J. Reedijk, *Eur. J. Inorg. Chem.* (2004) 4828-4835.
- [27] D. R. Hare, D. E. Wemmer, S.-H. Chow, G. Droby, B. R. Reid, *J. Mol. Biol.* **171** (1983) 319.
- [28] M. A. Akkerman, C. A. G. Haasnoot, C. W. Hilbers, *Eur. J. Biochem.* **173** (1988) 211.
- [29] M. A. Akkerman, E. W. Neijman, S. S. Wijmenga, C. A. G. Haasnoot, C. W. Hilbers, *J. Am. Chem. Soc.* **117** (1990) 7462.

- [30] S. S. Wijmenga, M. M. W. Mooren, C. W. Hilbers, in G. C. K. Roberts (Ed.): *NMR of Macromolecules: A Practical Approach*, Oxford University Press, New York 1993, p. 218-283.
- [31] D. Lemaire, M.-H. Fouchet, J. Kozelka, *J. Inorg. Biochem.* 53 (1994) 261-271.
- [32] F. J. M. van de Ven, C. W. Hilbers, *Eur. J. Biochem.* 178 (1988) 1-38.
- [33] S. S. Wijmenga, B. N. M. van Buuren, *Prog. Nucl. Mag. Res. Spect.* 32 (1998) 287-387.
- [34] A. T. M. Marcelis, C. G. van Kralingen, J. Reedijk, *J. Inorg. Biochem.* 13 (1980) 213-222.
- [35] A. T. M. Marcelis, C. Erkelens, J. Reedijk, *Inorg. Chim. Acta* 91 (1984) 129-135.
- [36] J. H. J. den Hartog, C. Altona, J. H. van Boom, A. T. M. Marcelis, L. J. Rinkel, G. Wille-Hazeleger, J. Reedijk, *Eur. J. Biochem.* 134 (1983) 485-495.
- [37] C. J. Van Garderen, C. Altona, J. Reedijk, *Inorg. Chem.* 29 (1990) 1481.
- [38] M. Polak, J. Plavec, A. Trifonova, a. Földesi, J. Chattopadhyaya, *J. Chem. Soc. Perkin Trans. I* (1999) 2835-2843.
- [39] C. Altona, M. Sundaralingam, *J. Am. Chem. Soc.* 94 (1972) 8205-8212.
- [40] D. Djuranovic, B. Hartmann, *J. Biomol. Struct. Dyn.* 20 (2003) 771-788.
- [41] R. E. Dickerson, M. Bansal, C. R. Calladine, S. Diekman, W. N. Hunter, O. Kennard, *J. Mol. Biol.* 205 (1989) 787-791.
- [42] R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* 6 (1989) 655-667.
- [43] P. Varnai, D. Djuranovic, R. Lavery, B. Hartmann, *Nucleic Acids Res.* 30 (2002) 5398-5406.
- [44] S. O. Ano, Z. Kuklenyik, L. G. Marzilli, in B. Lippert (Ed.): *Cisplatin: Chemistry and Biochemistry of a Leading Anticancer Drug*, Verlag Helvetica Chimica Acta, Zürich 1999, p. 247-291.
- [45] J. Kozelka, S. Archer, G. A. Petsko, S. J. Lippard, G. J. Quigley, *Biopolymers* 26 (1987) 1245-1271.
- [46] T. E. Cheatham III, P. Cieplak, P. A. Kollman, *J. Biomol. Struct. Dyn.* 16 (1999) 845-862.
- [47] P. Vichi, F. Coin, J.-P. Renaud, W. Vermeulen, J. H. J. Hoeijmakers, D. Moras, J.-M. Egly, *EMBO J.* 16 (1997) 7444-7456.

- [48] R. Visse, M. de Ruijter, J. Brouwer, J. A. Brandsma, P. van de Putte, *J. Biol. Chem.* 266 (1991) 7609-7617.
- [49] M. H. Lamers, A. Perrakis, J. H. Enzlin, H. H. K. Winterwerp, N. de Wind, T. K. Sixma, *Nature* 407 (2000) 711-717.
- [50] G. Obmolova, C. Ban, P. Hsieh, W. Yang, *Nature* 407 (2000) 703-710.
- [51] P. M. Pil, S. J. Lippard, *Science* 256 (1992) .
- [52] J. J. Love, X. Li, D. A. Case, K. Giese, R. Grosschedl, P. E. Wright, *Nature* 376 (1995) 791-795.
- [53] M. H. Werner, J. R. Huth, A. Gronenborn, G. M. Clore, *Cell* 81 (1995) 705-714.
- [54] S. Komeda, S. Bombard, S. Perrier, J. Reedijk, J. Kozelka, *J. Inorg. Biochem.* 96 (2003) 357-366.
- [55] V. Sklenar, M. Piotto, R. Leppik, V. Saudek, *J. Magn. Res.* 102 (1993) 241.
- [56] P. J. Kraulis, *J. Magn. Res.* 24 (1989) 627.
- [57] D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, Jr., R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, R. J. Radmer, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. K. Weiner, P. A. Kollman, , University of California, San Francisco 1999.
- [58] T. A. Darden, D. York, L. G. Pedersen, *J. Chem. Phys.* 98 (1993) 10089-10092.
- [59] U. Essman, L. Perera, M. Berkowitz, T. A. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* 103 (1995) 8577-8593.
- [60] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* 23 (1977) 327-341.
- [61] M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. A. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, J. A. Pople, , Gaussian, Inc., Pittsburgh, PA, USA 1995.
- [62] P. J. Hay, W. R. Wadt, *J. Chem. Phys.* 82 (1985) 299-310.
- [63] S. Arnott, P. Campbell-Smith, P. Chandrasekharan, *CRC Handbook of Biochemistry*, Vol. 2 1976.

- [64] A. Dalke, W. Humphrey, J. Ulrich, , Theoretical Biophysics Group, University of Illinois and Beckman Institute, Urbana, USA 1997.

Supporting Material

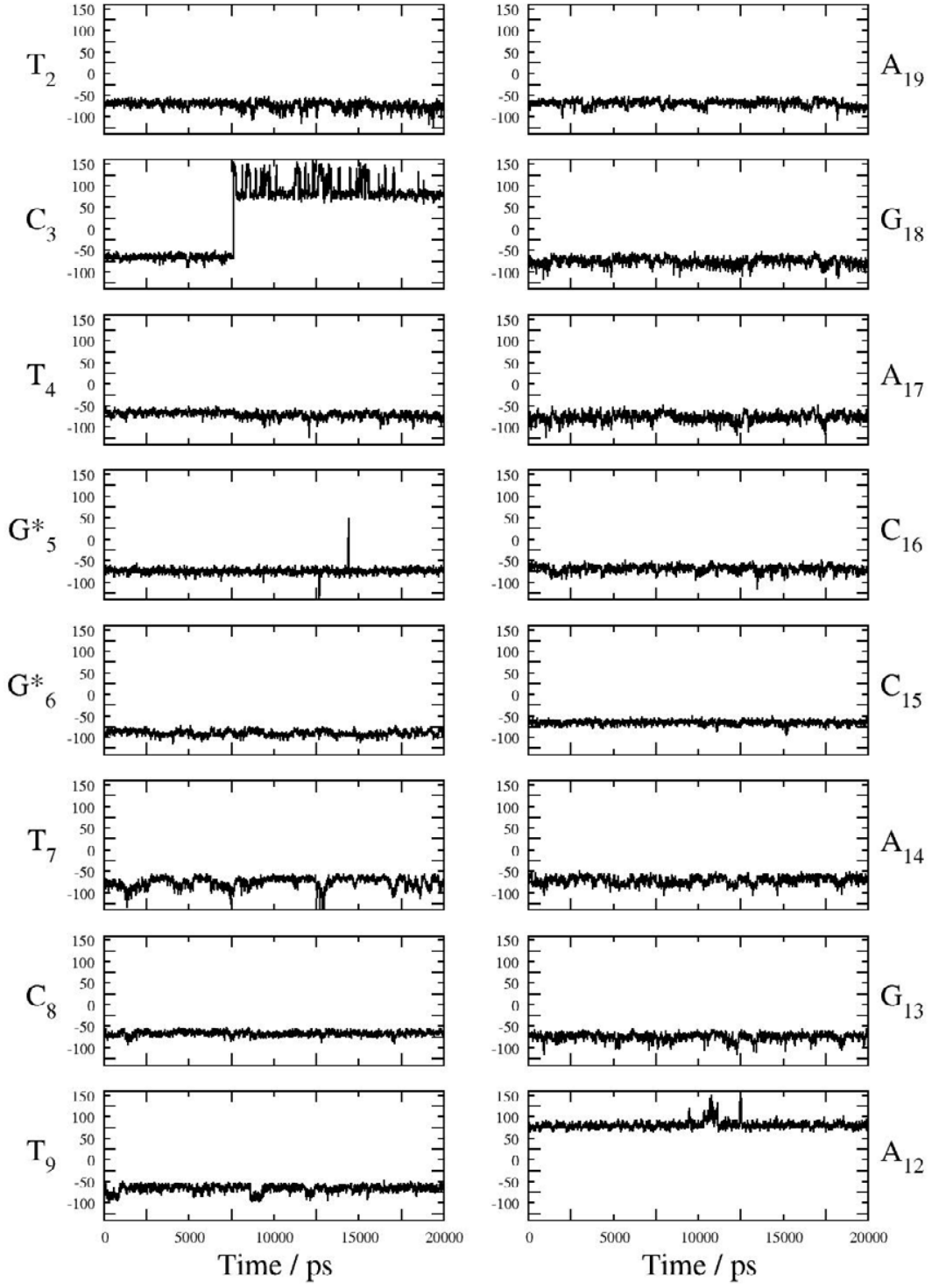
to

A pyrazolato-bridged dinuclear platinum(II) complex induces only minor distortions upon DNA-binding

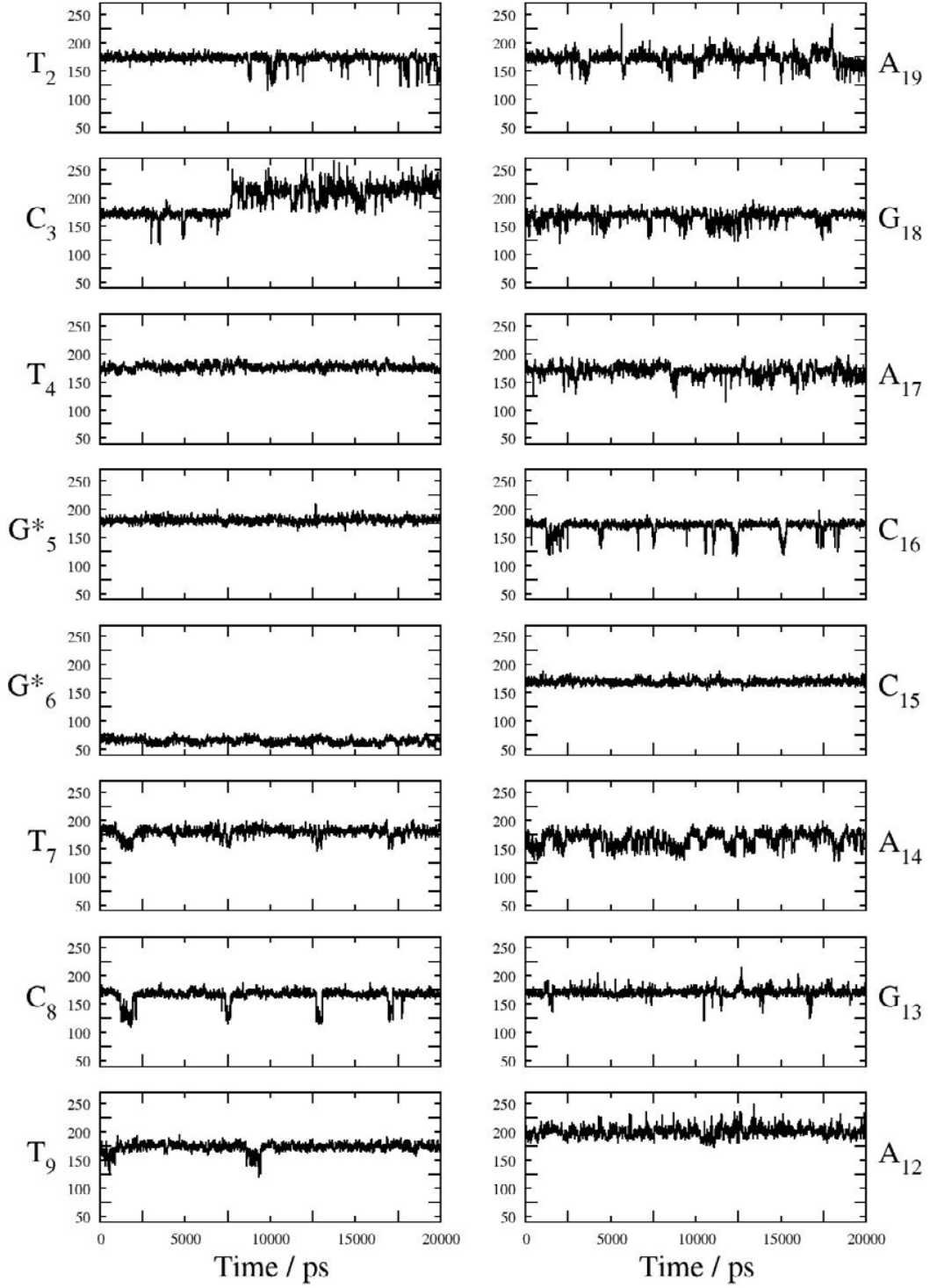
Stéphane Teletchéa,^[a,b] Seiji Komeda,^[a,c,d] Jan-Maarten Teuben,^[c,e] Miguel-Angel Elizondo-Riojas,^[b,f] Jan Reedijk,^{[c]*} and Jiří Kozelka^{[b]*}

Figure S1 (23 pages). Evolution of all backbone torsion angles and helicoidal parameters of dsII (excluding the terminal base pairs) during the MD simulation.

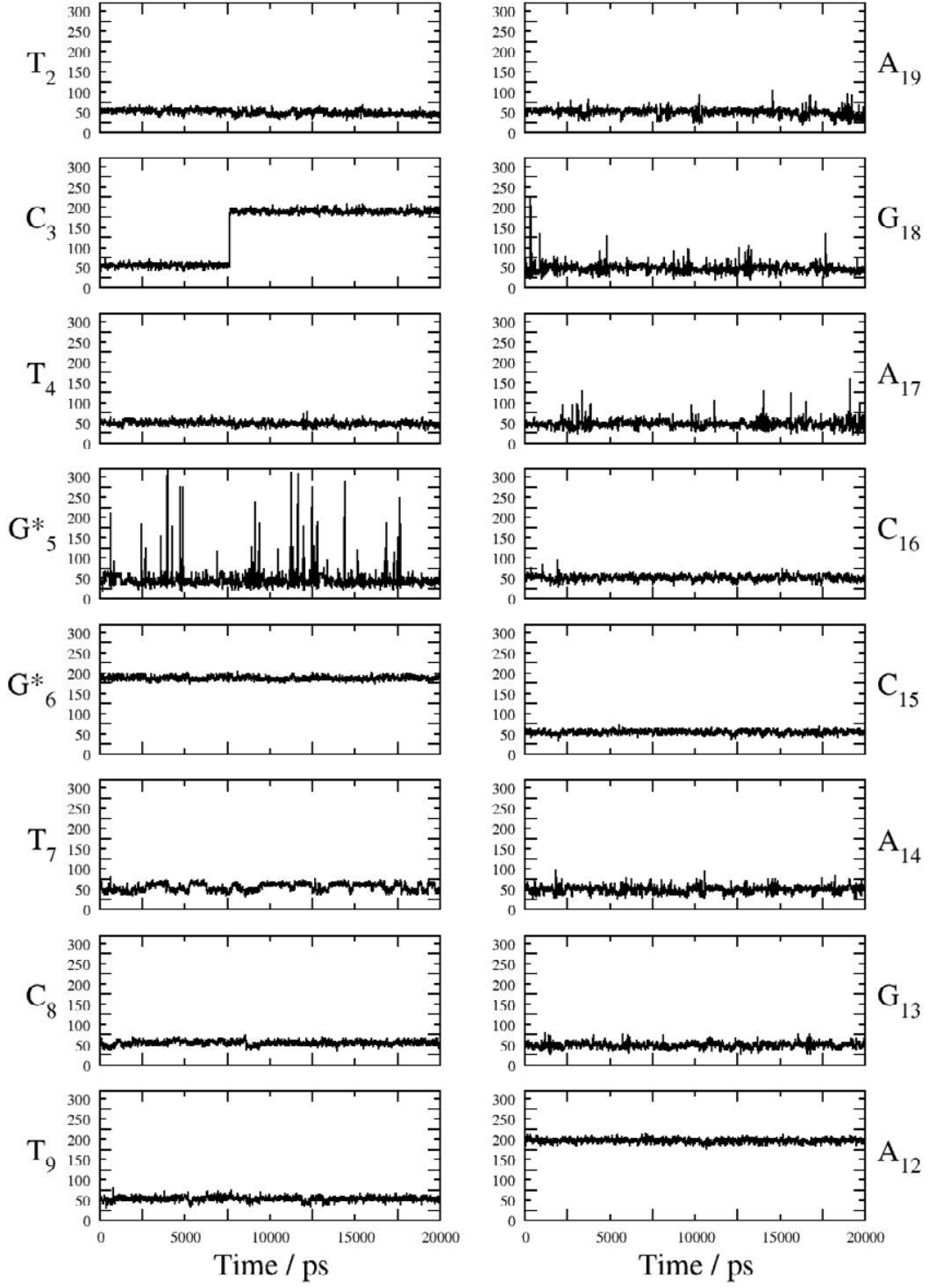
- alpha -



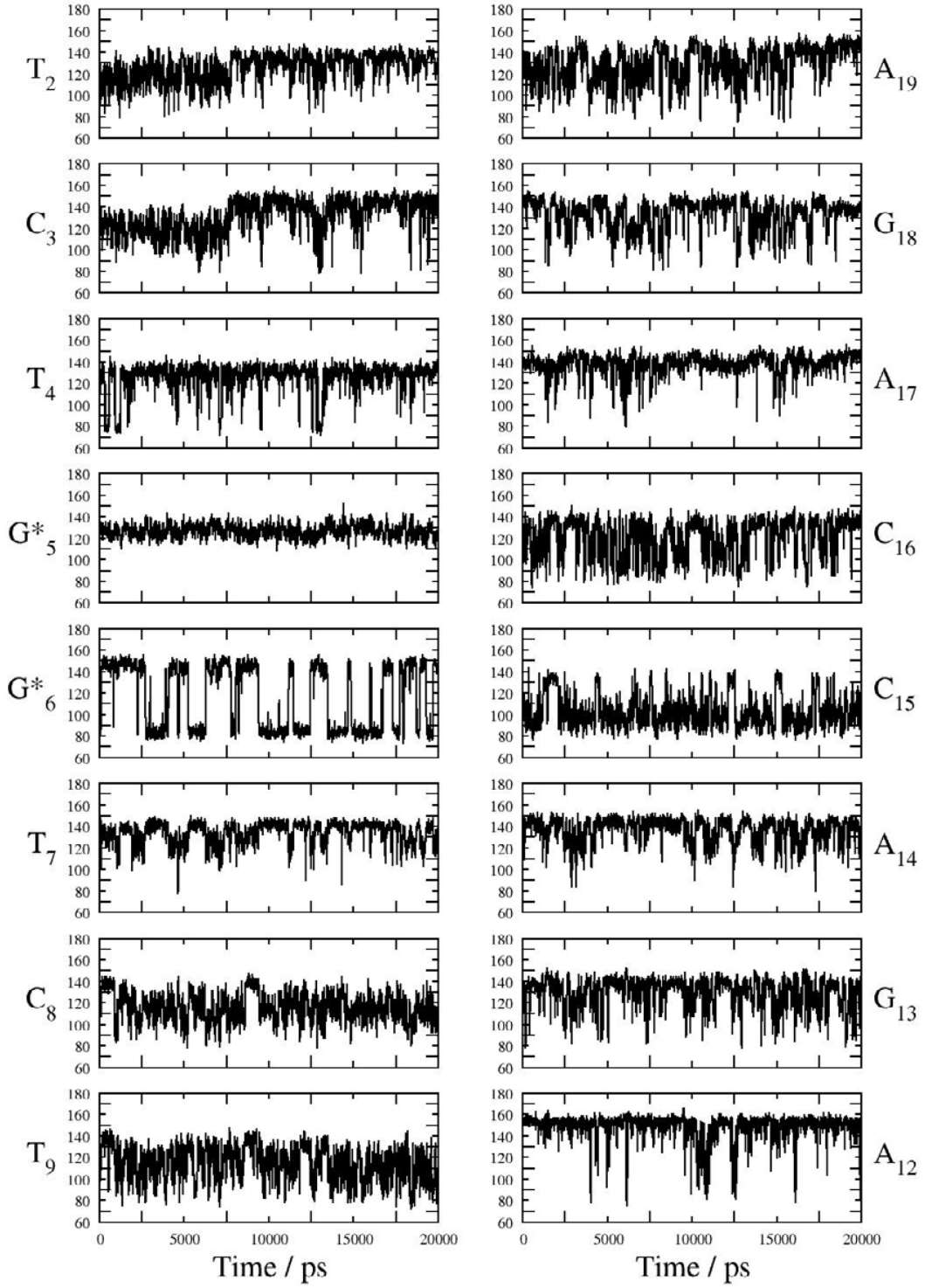
- beta -



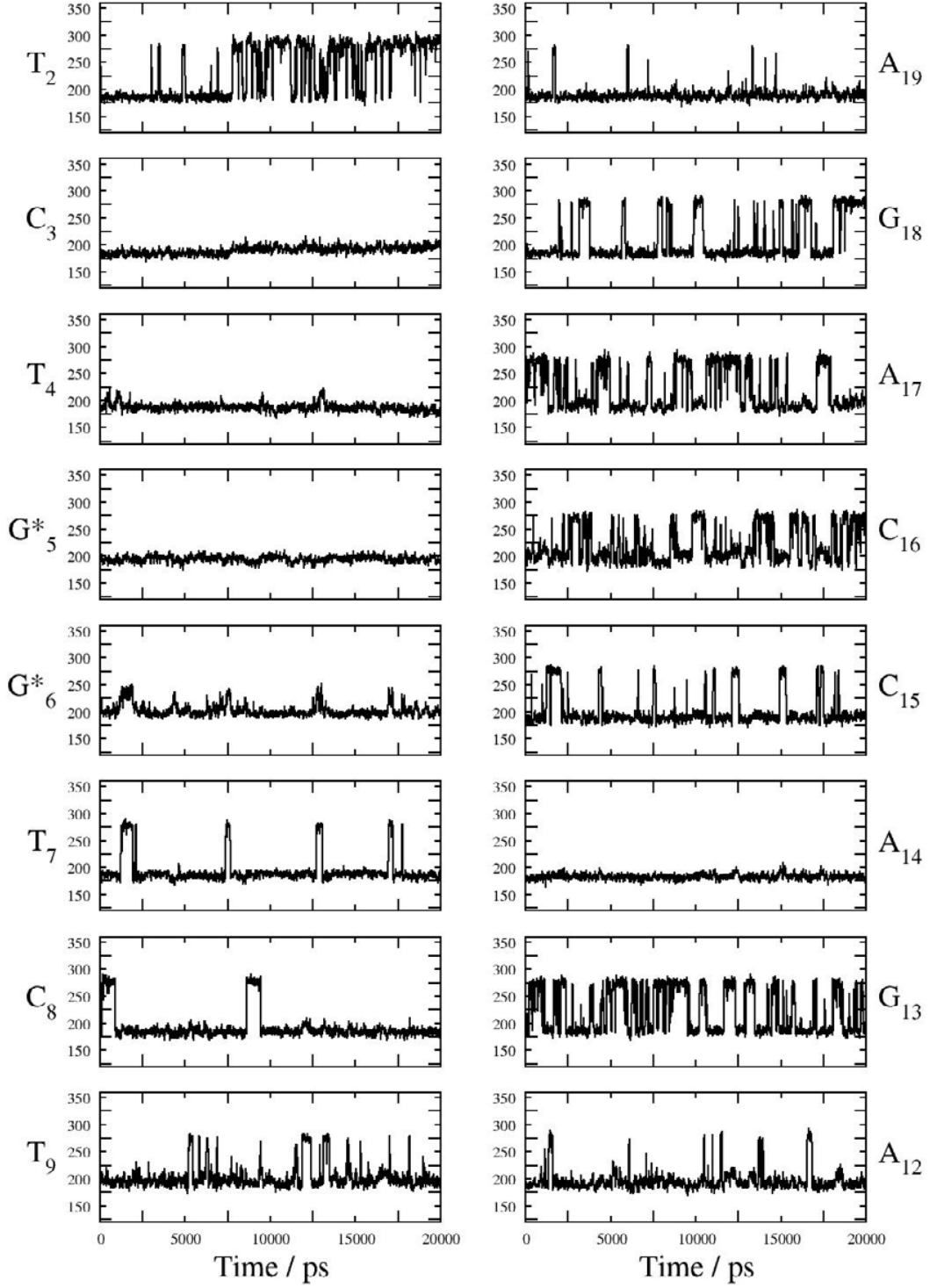
- gamma -



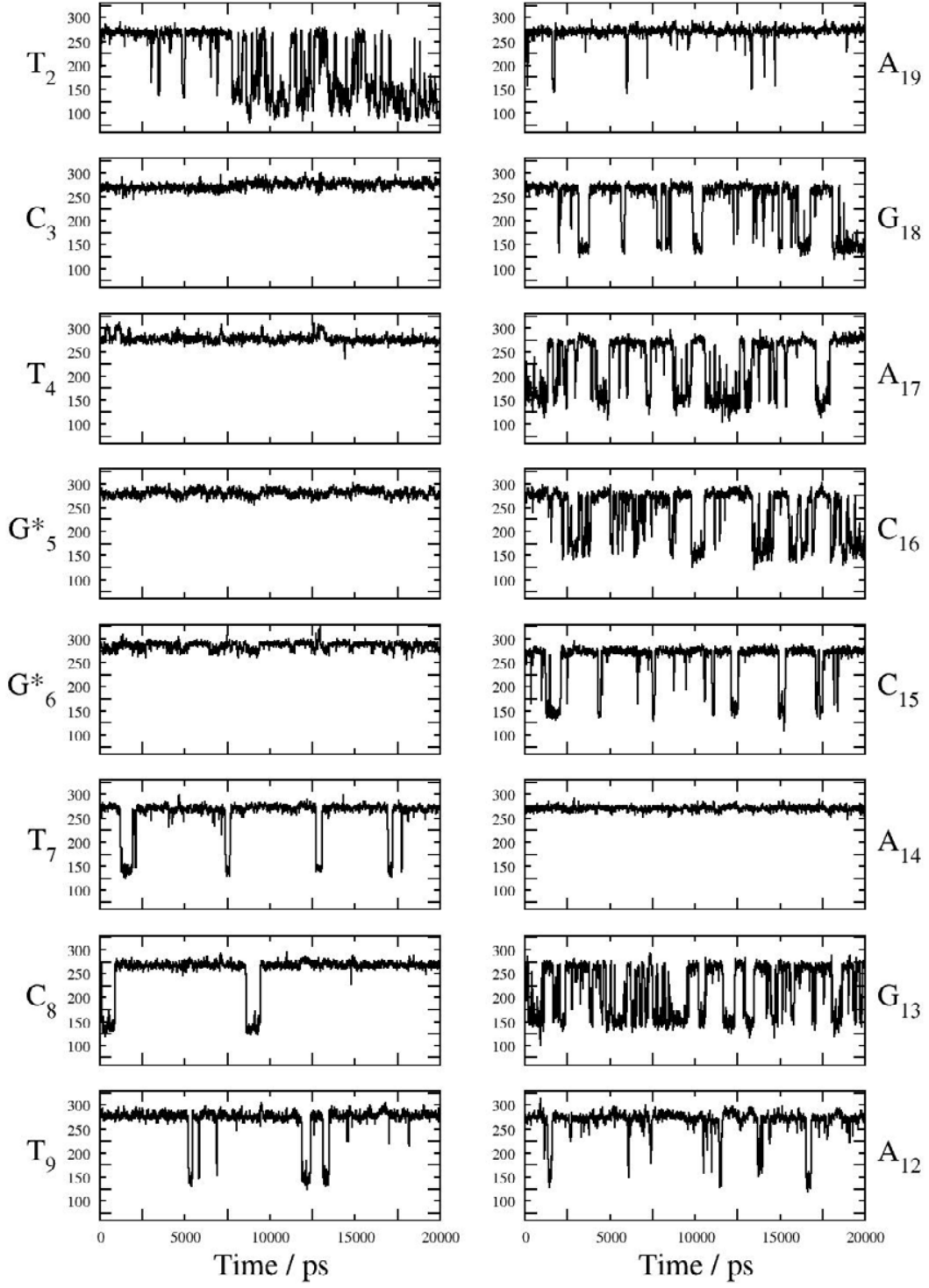
- delta -



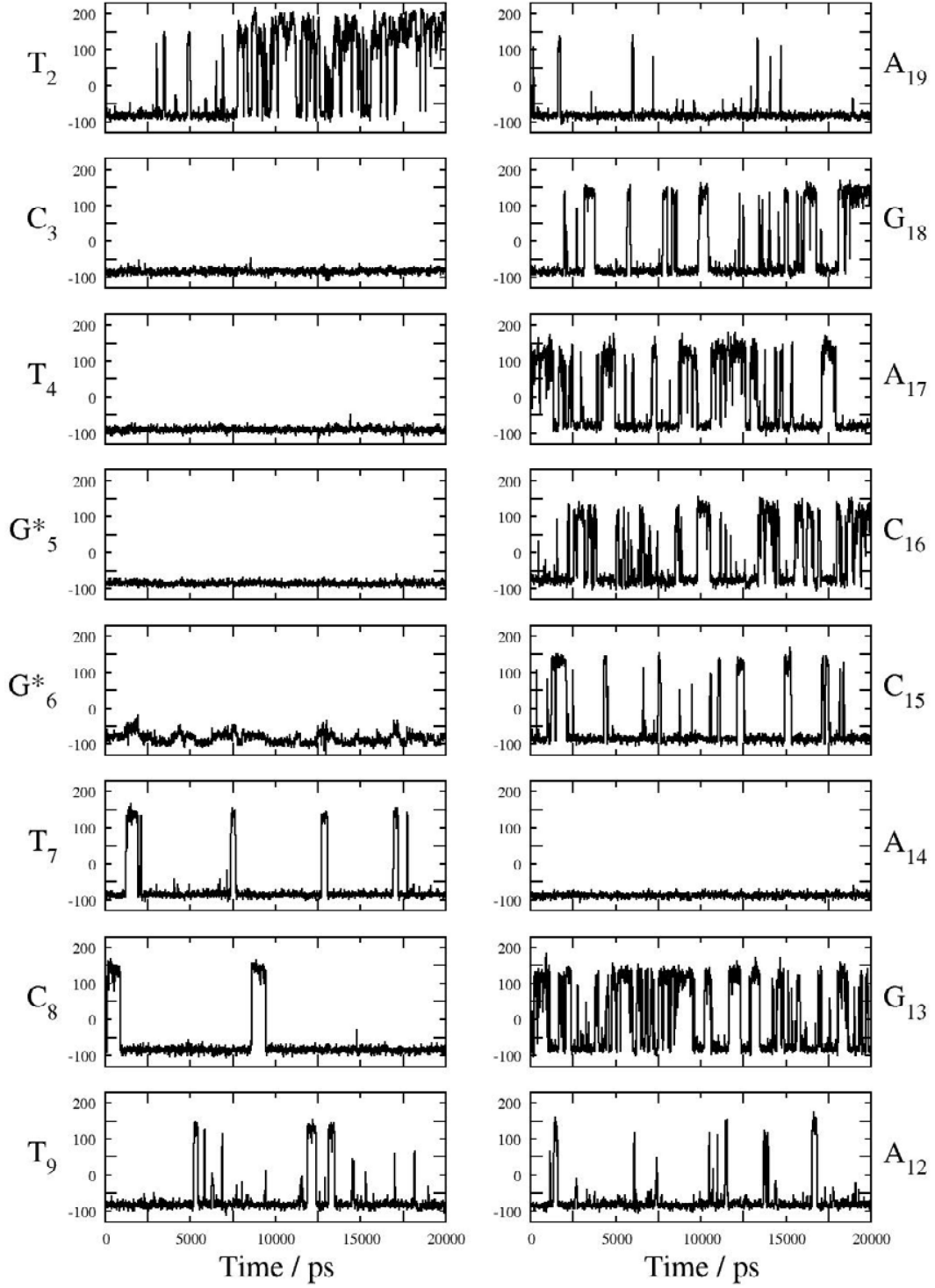
- epsilon -



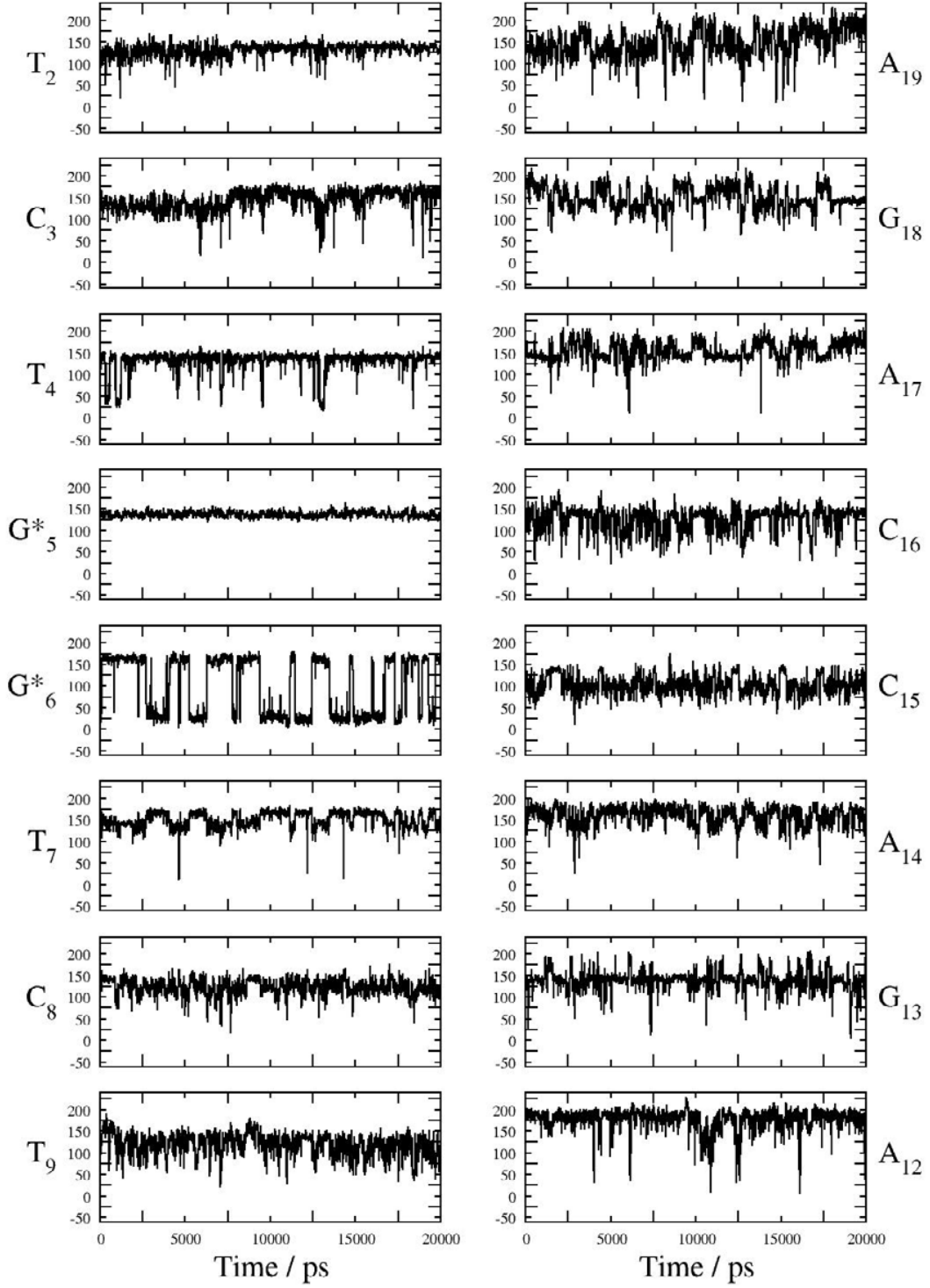
- zeta -



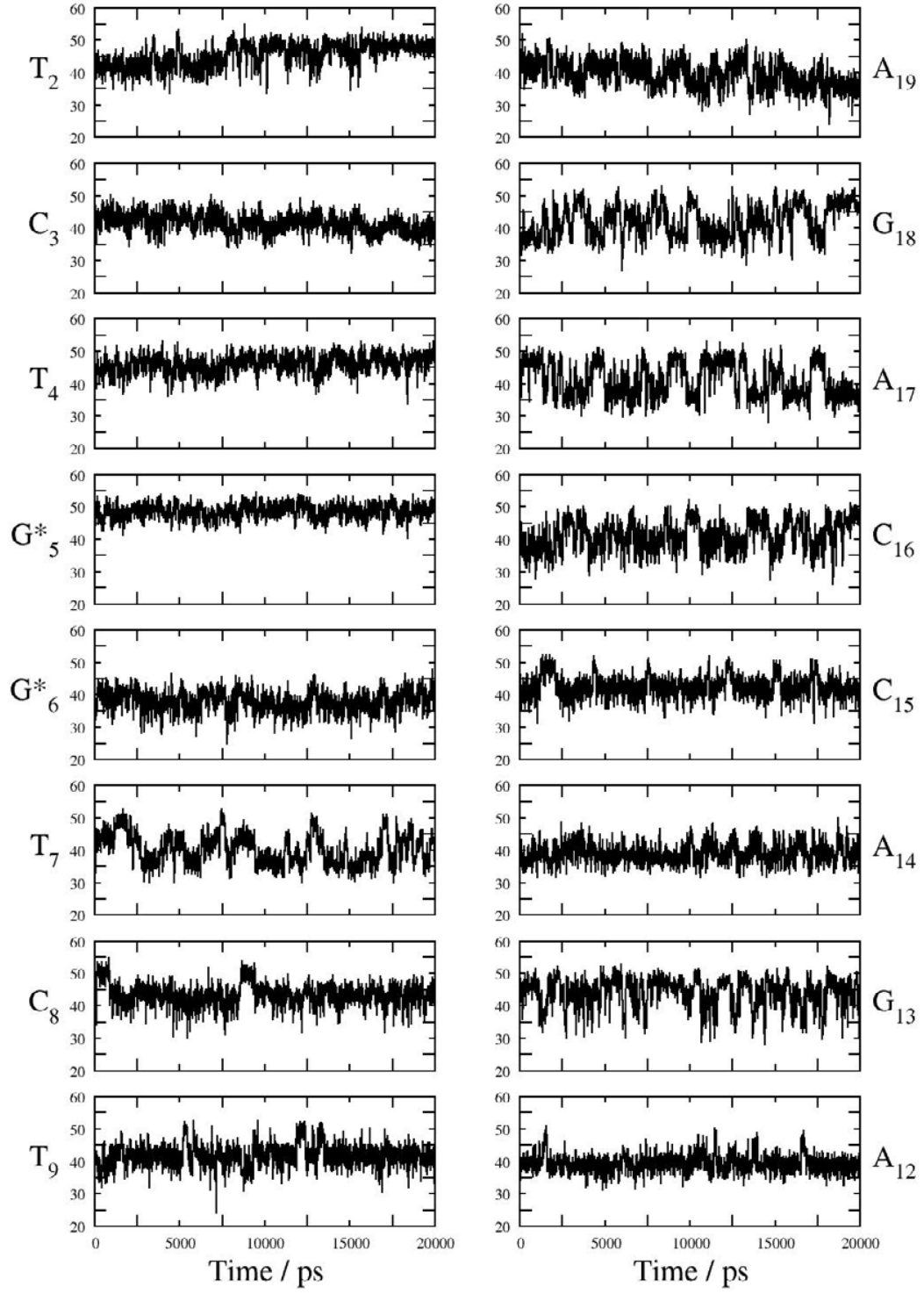
- epmoinsze -



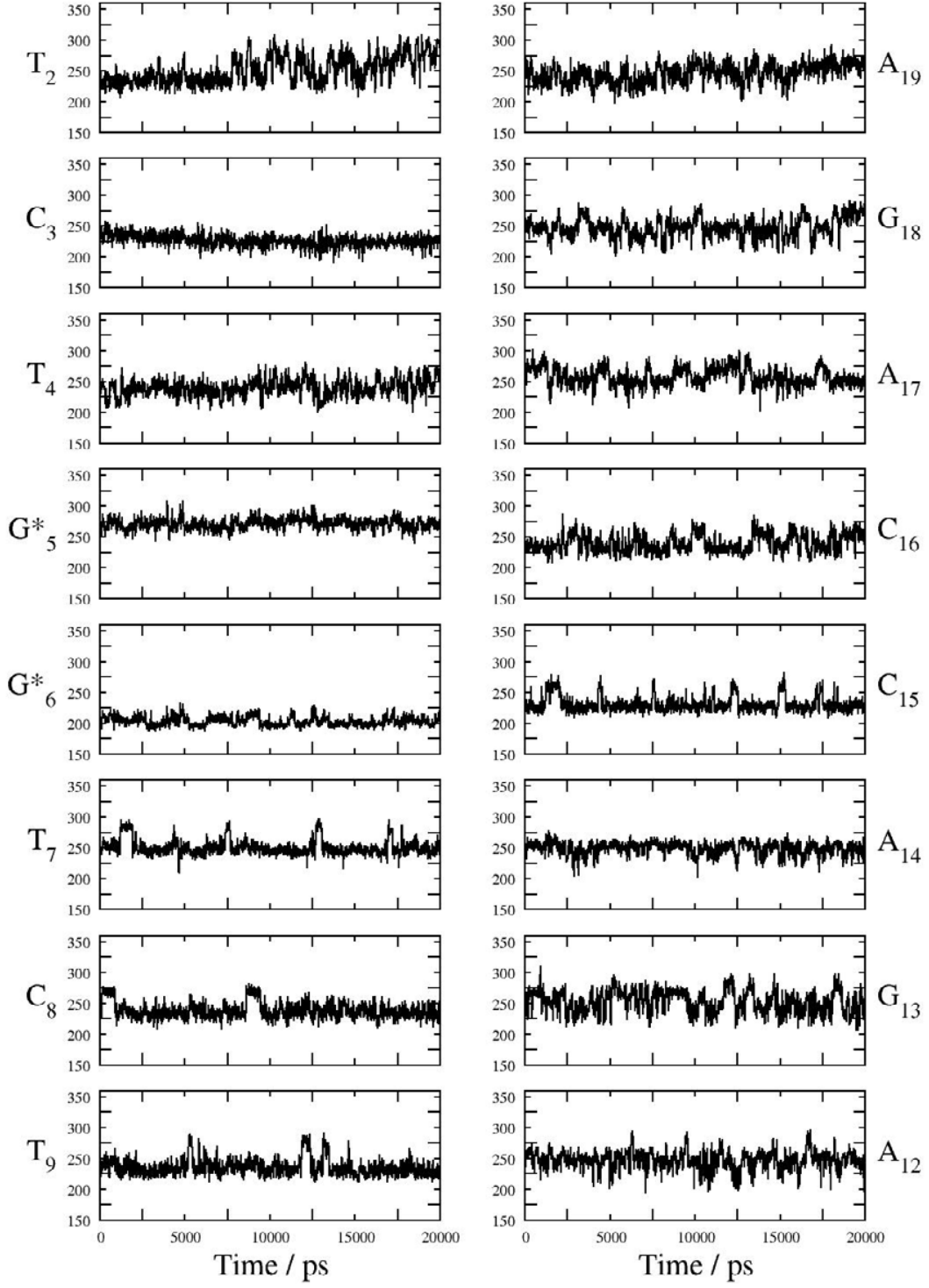
- Phase -



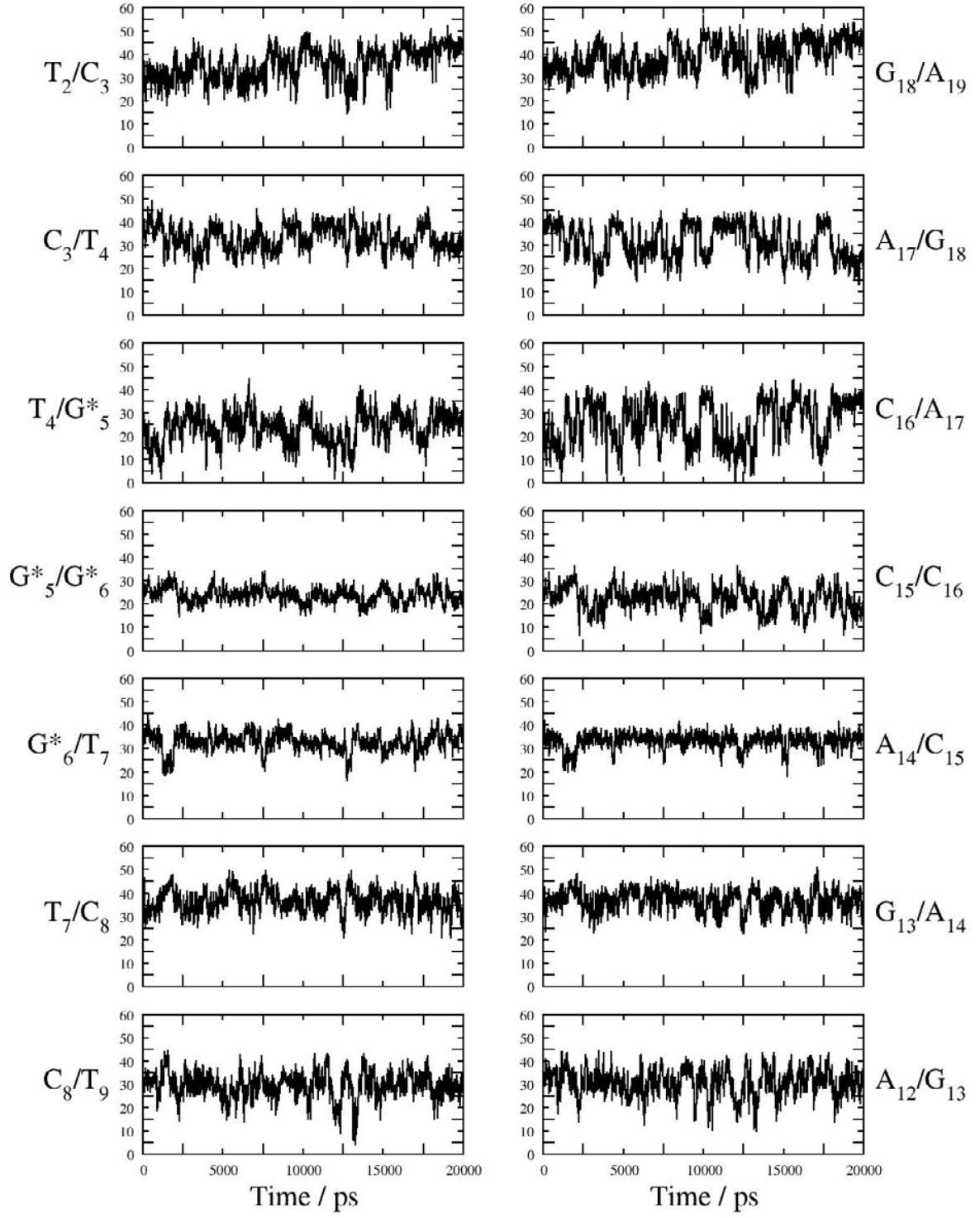
- Amplitude -



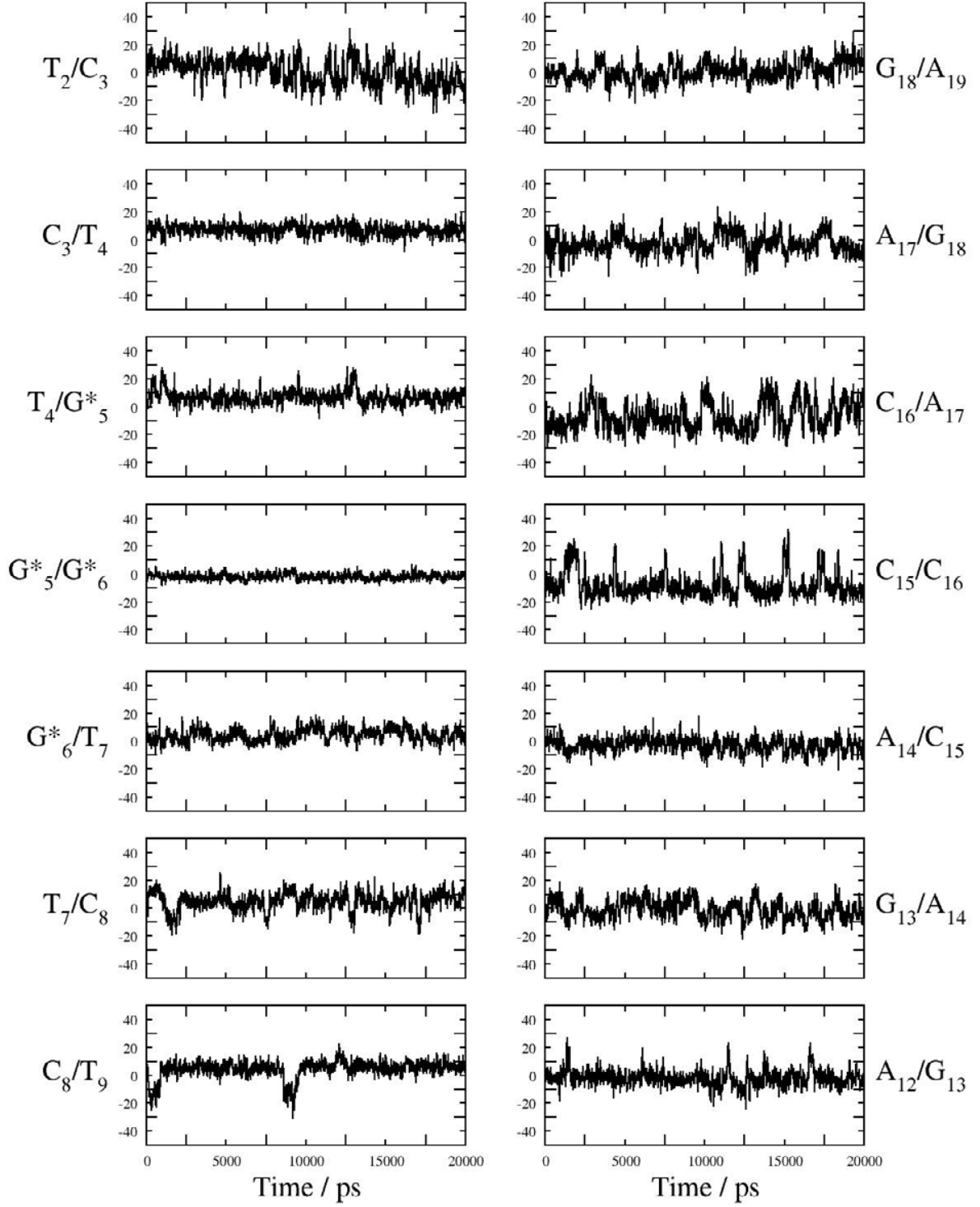
- chi -



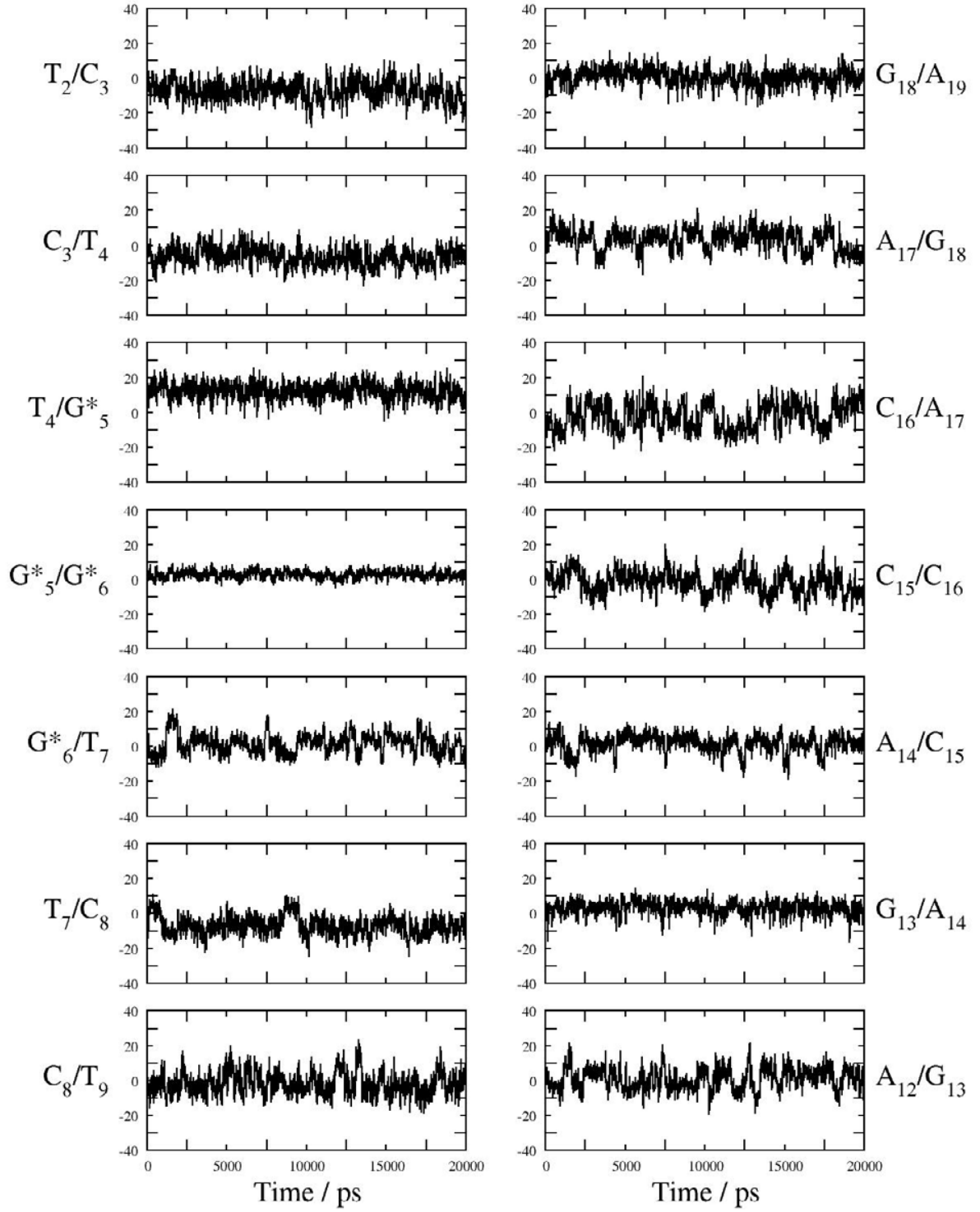
- Twist_Omega -



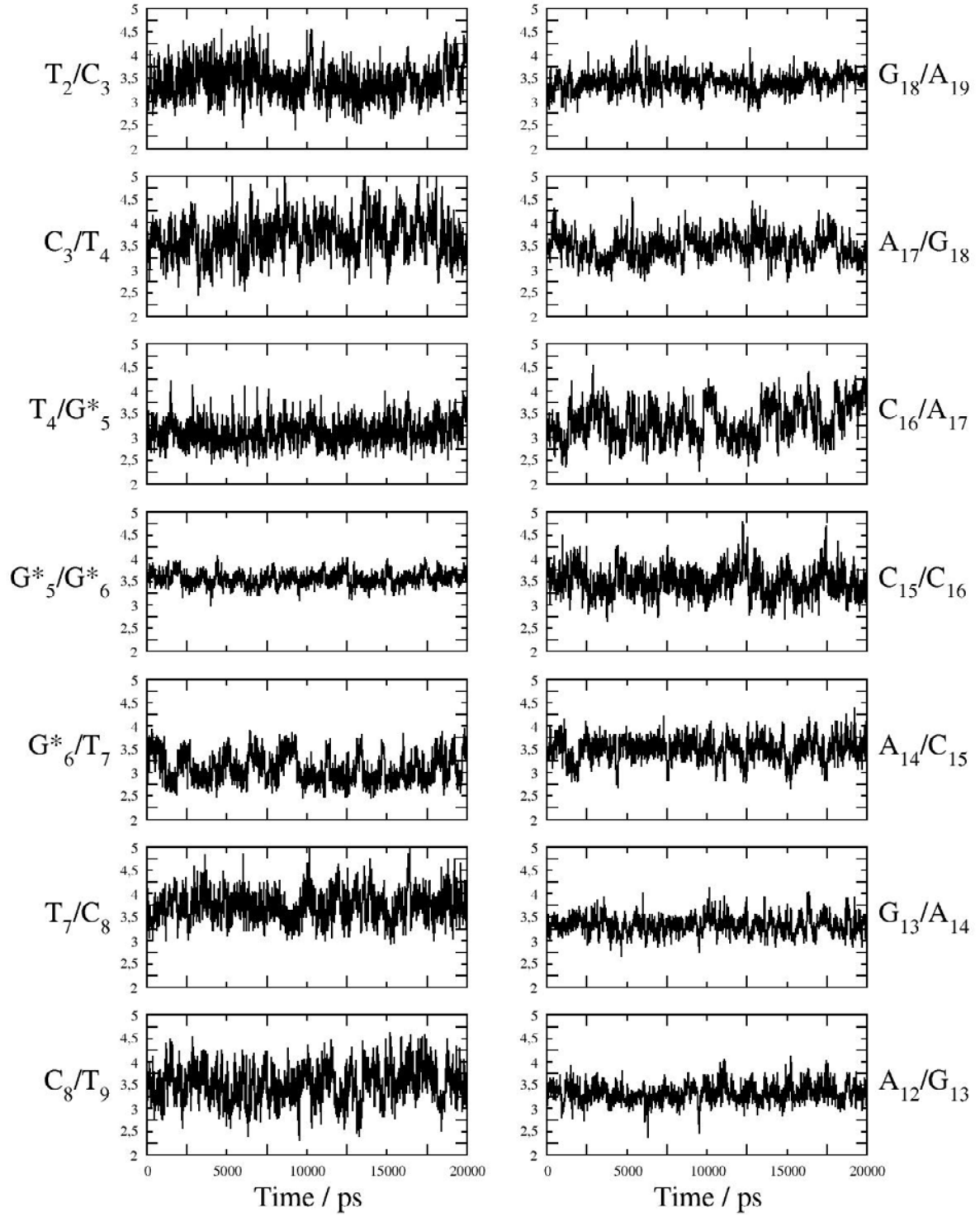
- Roll_rho -



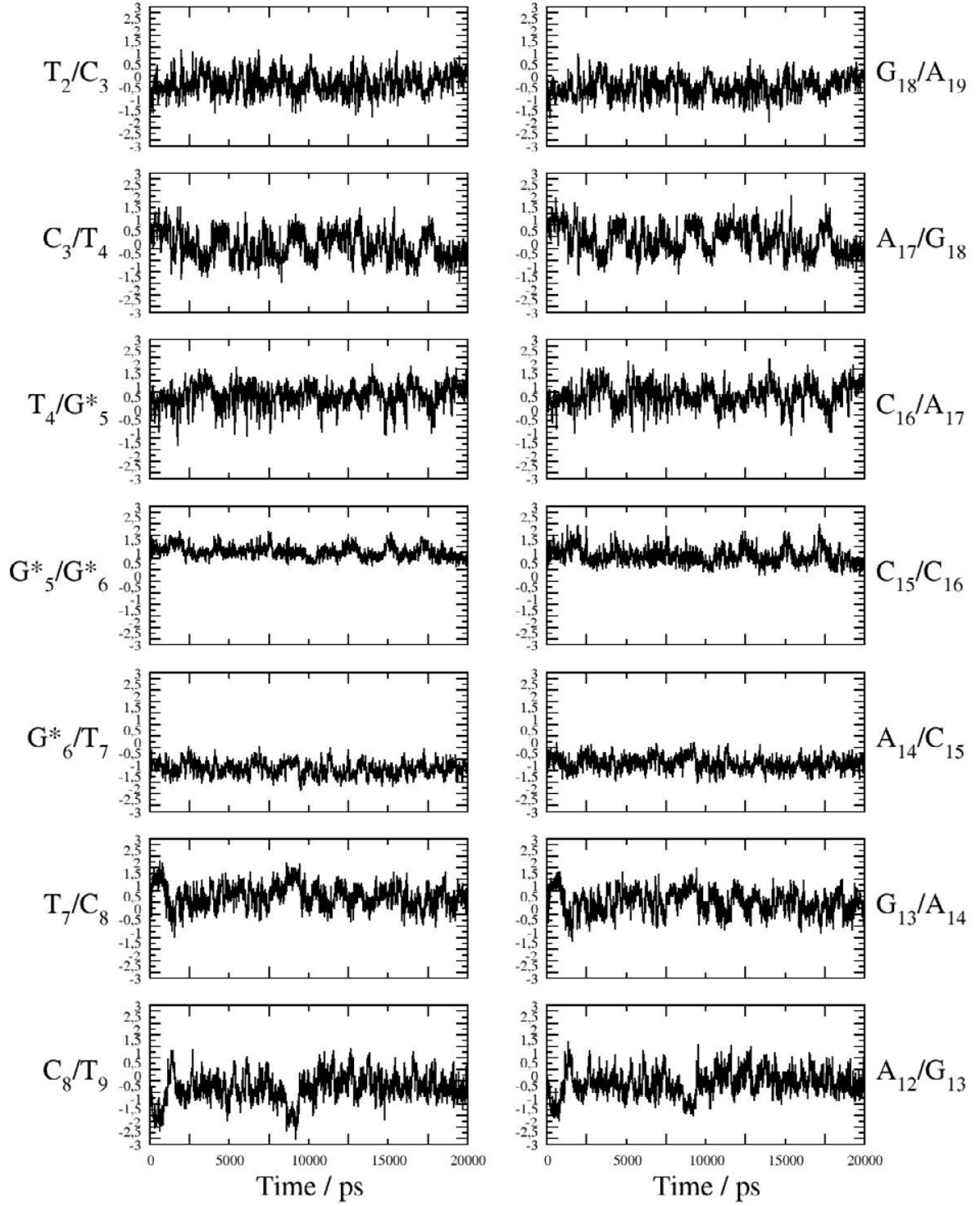
- Tilt_tau -



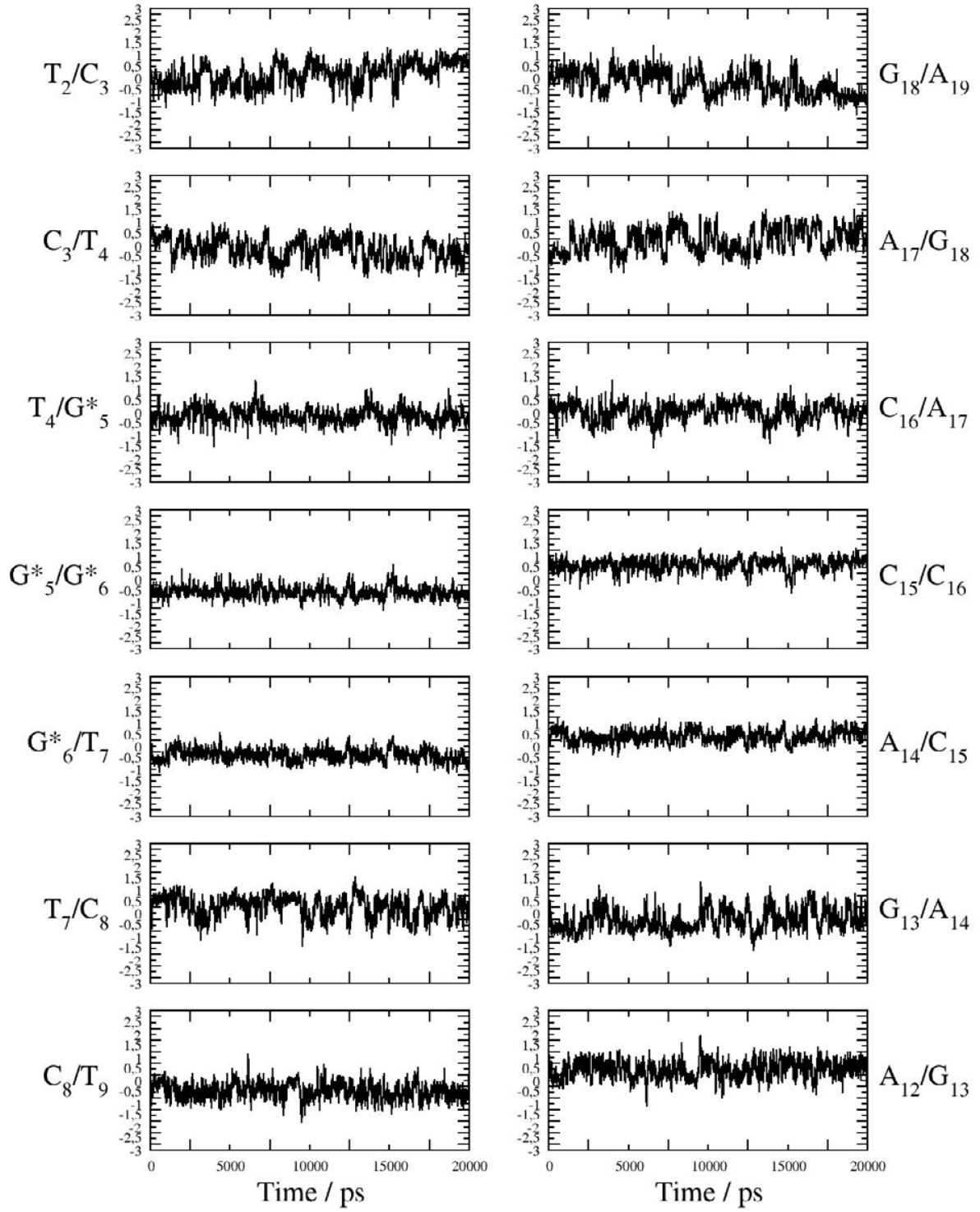
- Rise_Dz -

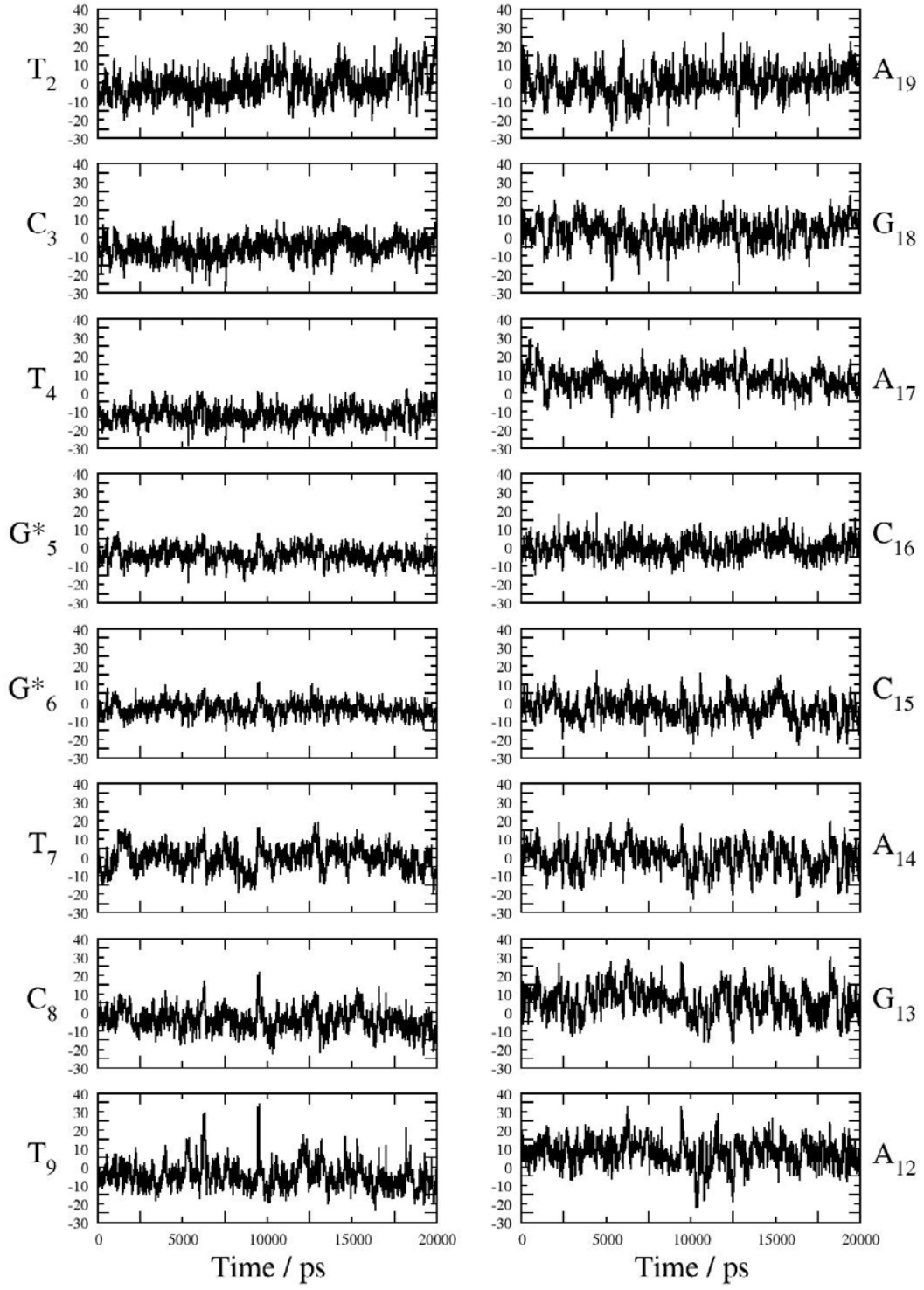


- Shift_Dx -

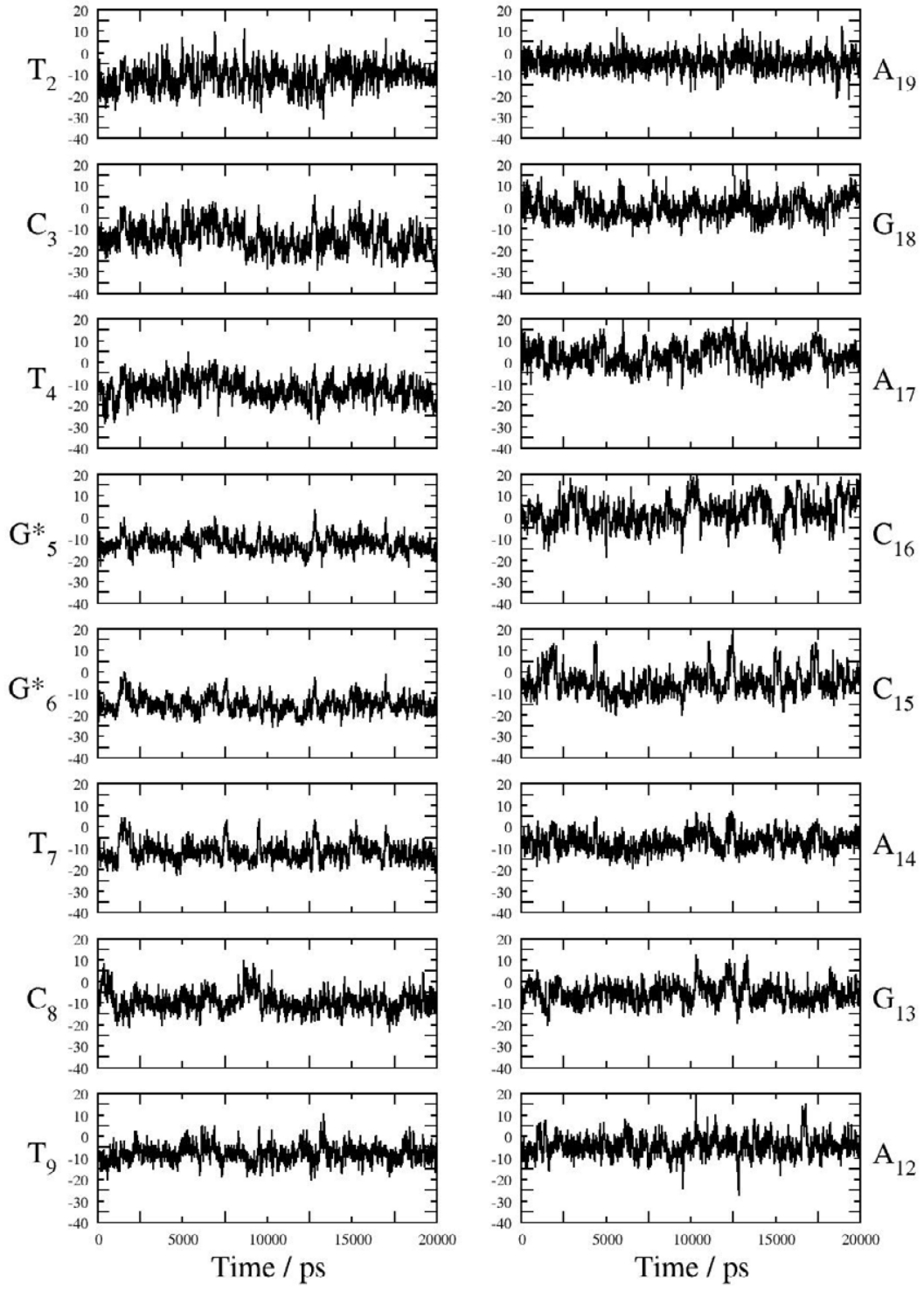


- Slide_Dy -

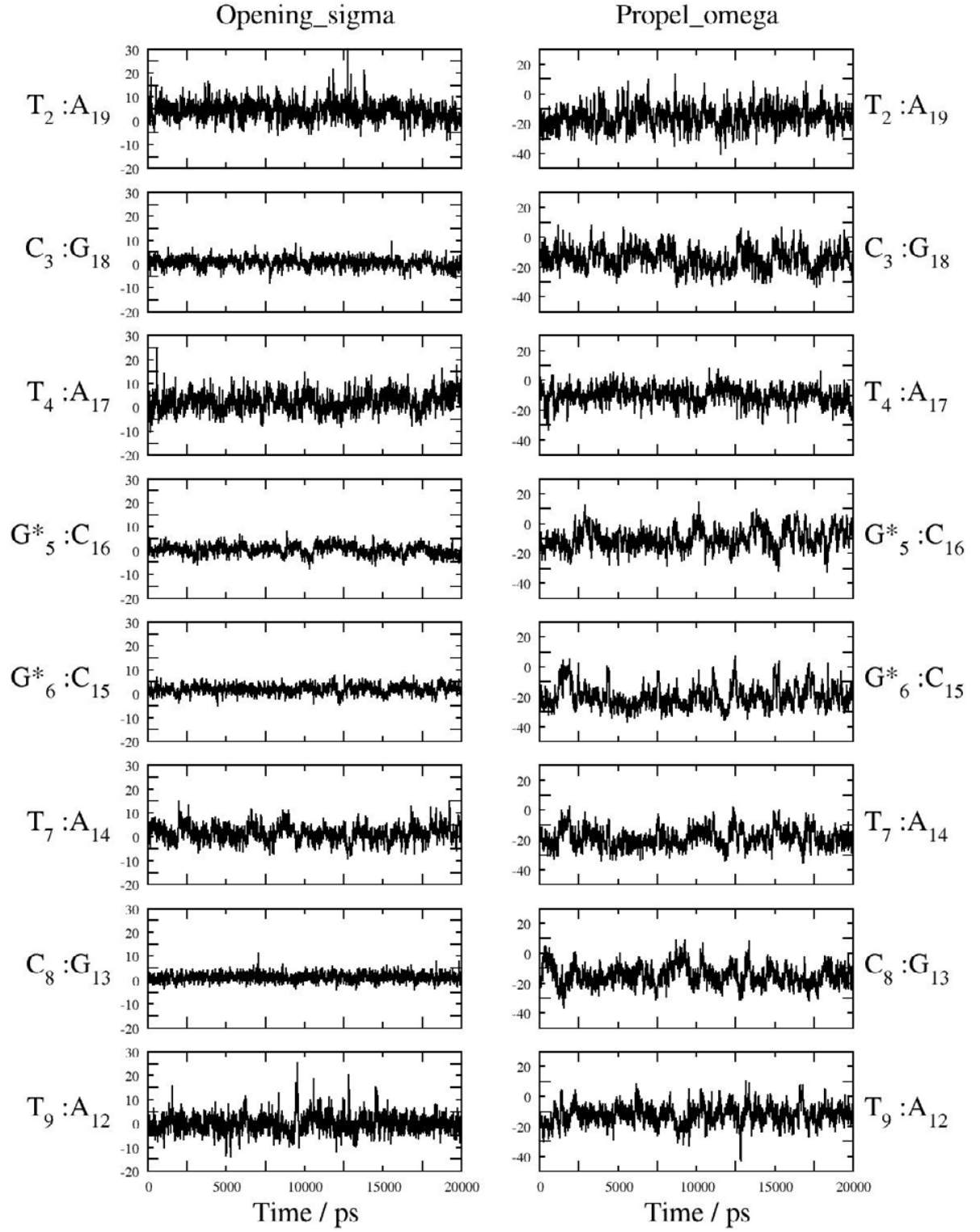


- Incl_{in}_eta -

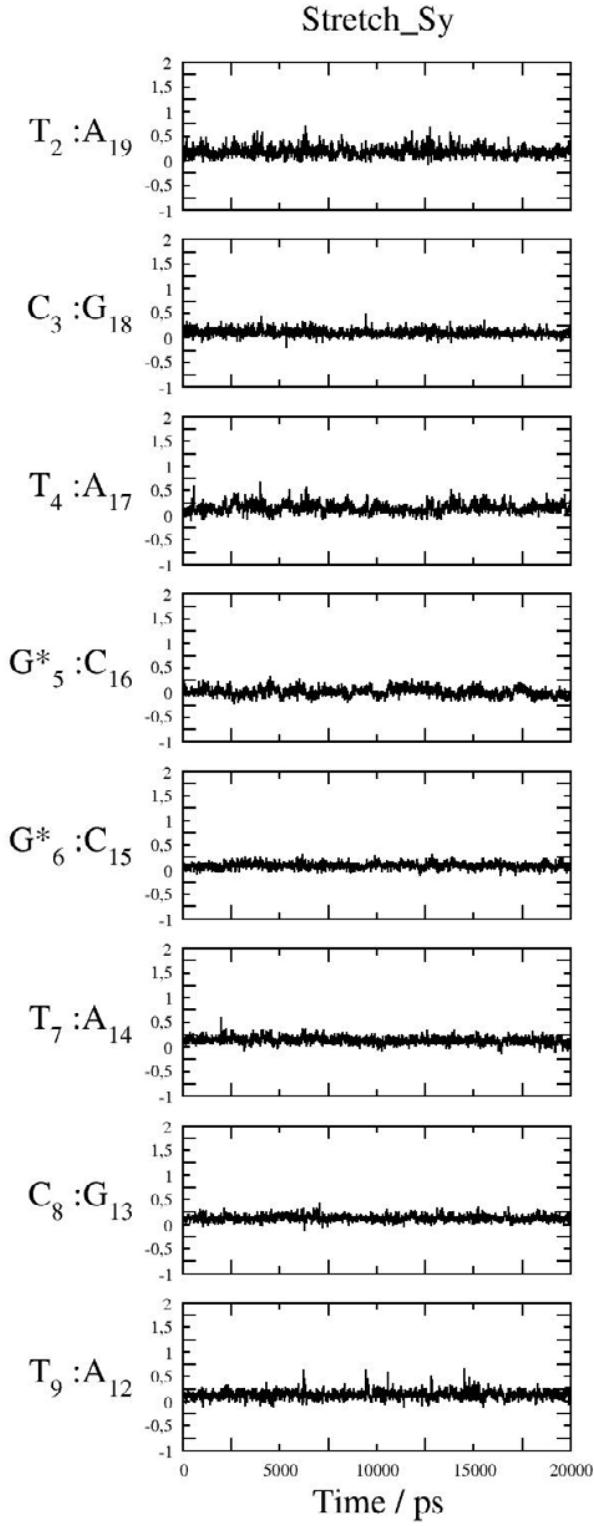
- Tip_theta -



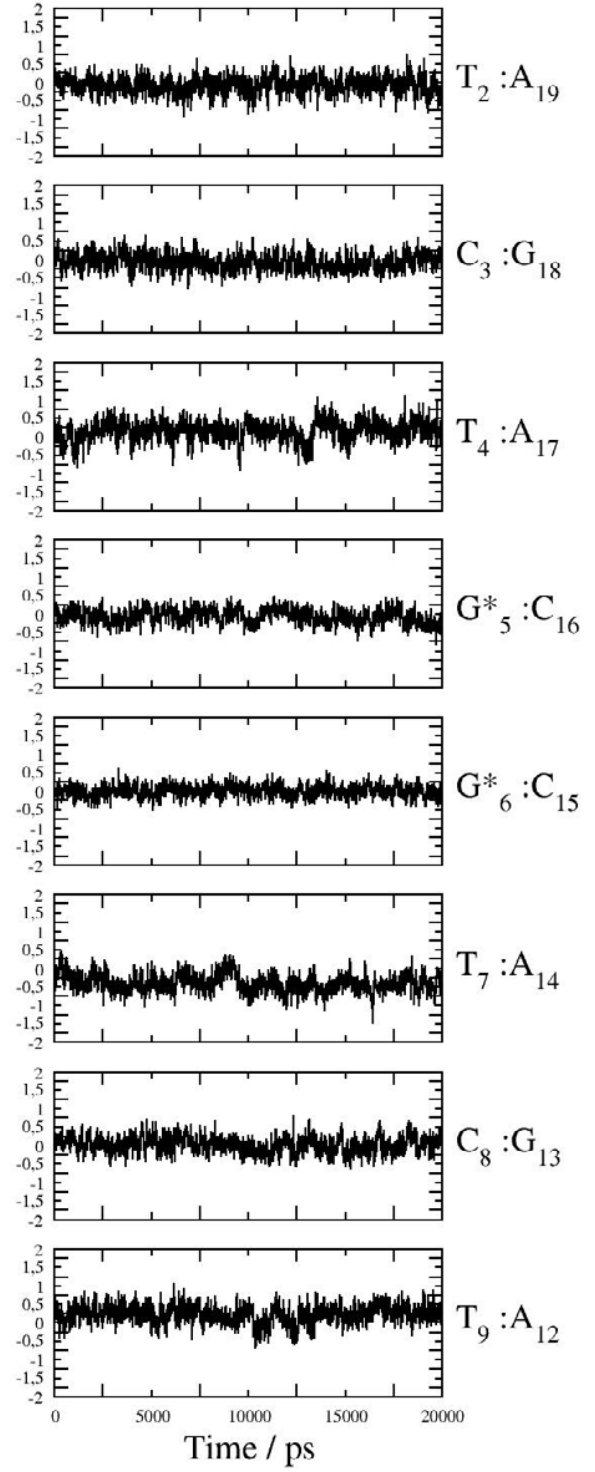
- Opening_sigma -



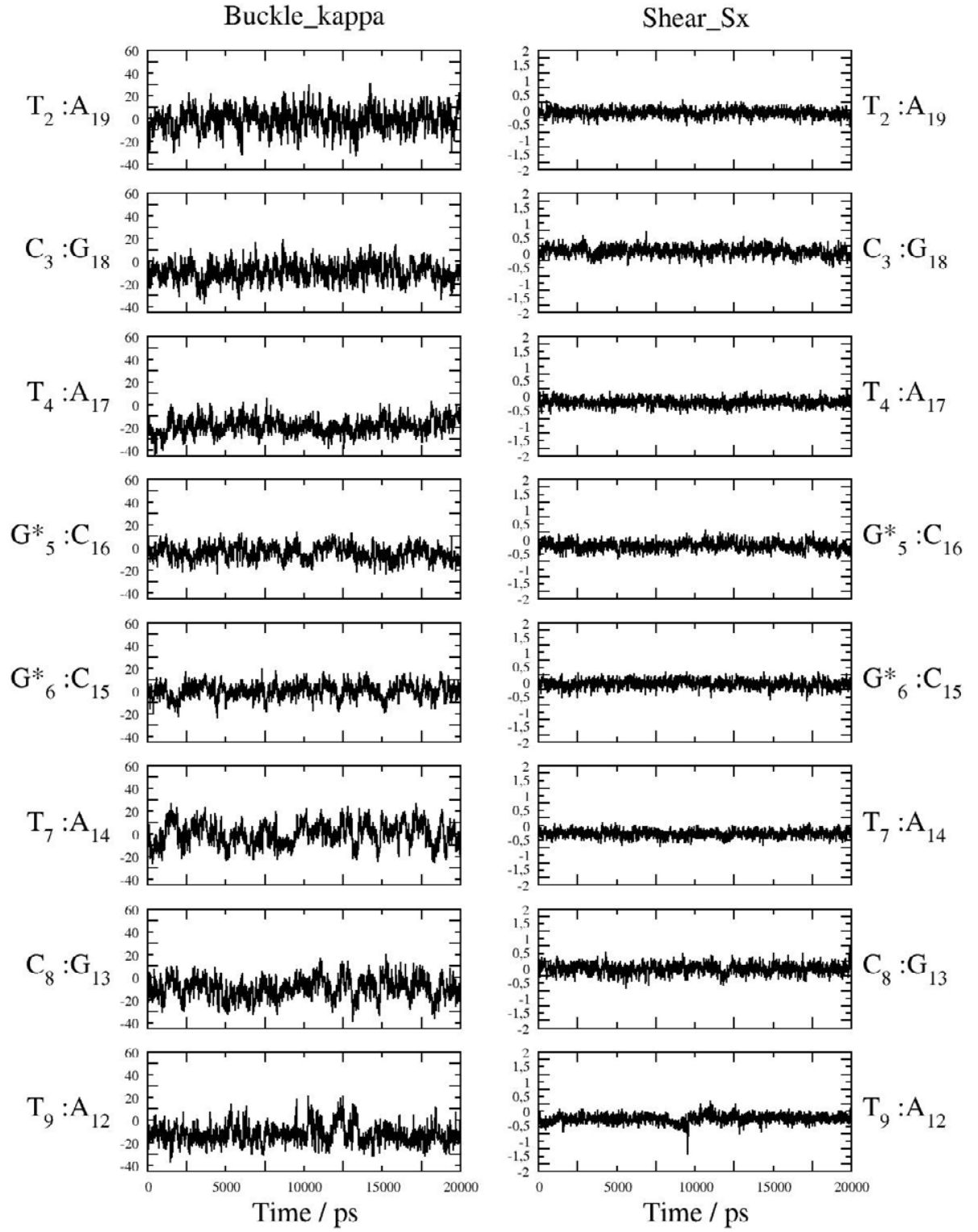
- Stretch_Sy -



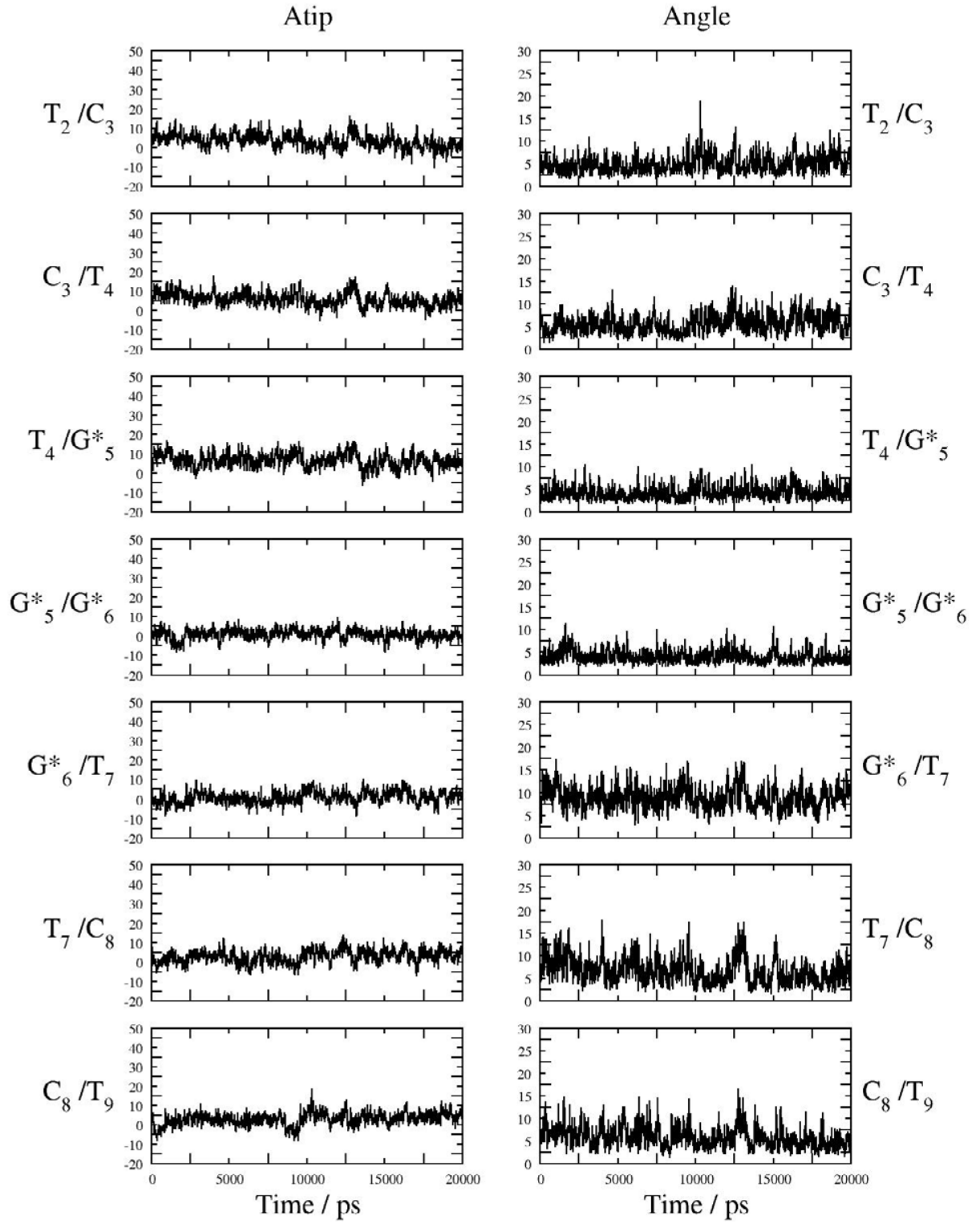
Stagger_Sz



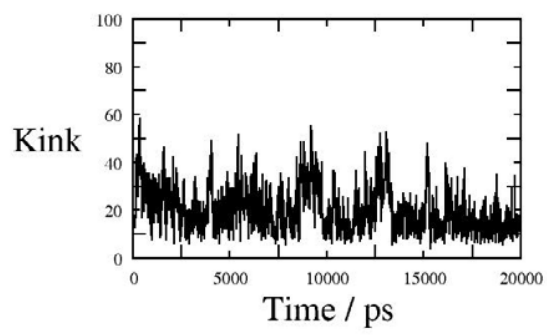
- Buckle_kappa -



- Atip -



- Kink -



Cinquième partie

Les composés platinés dans AMBER

Nous avons vu dans l'introduction comment est représenté chaque atome dans le modèle « tout atome » utilisé dans la suite logicele AMBER. Les monomères des macromolécules biologiques courantes (acides aminés, nucléotides), quelques solvants (méthanol, éthanol, eau), quelques hétéroatomes (brome, chlore, fer, iode, sodium, ...) sont définis en standard.

Quand des molécules comportent des atomes non définis comme c'est le cas pour les composés platinés, il faut ajouter leur description au champ de force.

Chapitre 1

Caractéristiques chimiques du cisplatine

L'élément chimique platine est un métal de transition qui comporte 78 électrons. Dans le degré d'oxydation II, 8 électrons sont disponibles sur sa couche externe de valence, il fait partie de la catégorie des métaux 'd8'. Le platine II adopte alors une géométrie plan-carré, ses quatre ligands étant situés dans le même plan, comme illustré pour le cisplatine par la figure 1.1.

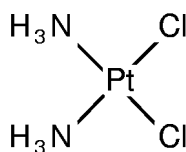


FIG. 1.1 – Le cisplatine

Le cisplatine réagit avec les groupements nucléophiles et se lie préférentiellement aux atomes d'azote et de soufre. Les deux ligands ammines en position *cis* sont inertes à la substitution. Les deux chlorures sont relativement labiles et peuvent être substitués en solution par des ligands aqua, la charge globale du complexe ainsi formé devient 2+. Dans la cellule, le cisplatine se lie principalement aux atomes N_7 des guanines de l'ADN du noyau et de la mitochondrie.

Chapitre 2

Adaptation du champ de force parm

98

Les composés de platine II que nous étudions dans le laboratoire présentent une géométrie plan-carré qui n'existe pas dans la définition standard des atomes dans le champ de force parm 98 [29]. L'ensemble du résidu platiné a dû être défini et intégré dans le champ de force, les paragraphes suivants présentent le détail des caractéristiques introduites dans le cadre du cisplatine ainsi que l'origine des données utilisées. La présentation exhaustive des fichiers utilisés est disponible en annexe, à la page 257.

Ce travail ayant été réalisé avant mon arrivée au laboratoire, je me suis attaché à valider cette paramétrisation puis à l'améliorer. Le cisplatine induit des déformations mécaniques importantes sur les guanines qu'il pontage, ces déformations ont fait l'objet d'études poussées [108, 60], ce qui a mené à l'utilisation d'angles impropres pour en rendre compte. L'effet inductif du platine sur les charges des guanines a été inclus à la suite de calculs de chimie quantique [60, 106, 108]. Les effets électroniques ont été alors limités à l'ensemble guanine-platine-guanine mais nous verrons plus bas qu'une paramétrisation spécifique des sucres et du squelette phosphodiester sera aussi nécessaire, notamment à la lecture des travaux de Plavec et ses collaborateurs [151].

La paramétrisation des pontages G-Pt-G dans le champ de force a été réalisée à partir de données structurales publiées [118, 170], extraites de la base de données du « *Cambridge Crystallographic Data Centre* » [92].

La figure 2.1 présente nos conventions de nommage des résidus adoptées pour l'incorporation du cis-Pt(NH₃)₂²⁺ dans le champ de force. Les atomes d'azote des deux ammines sont nommées NS et NW, la guanine située en 5' de l'adduit est notée GDA et est reliée au cis-Pt(NH₃)₂²⁺ par l'atome identifié comme NE, en 3' il s'agit respectivement de GDB et de NF.

Le tableau 2.1 détaille les paramètres utilisés lors des simulations de dynamique moléculaire avec le champ de force parm 98 et la suite logicielle AMBER.

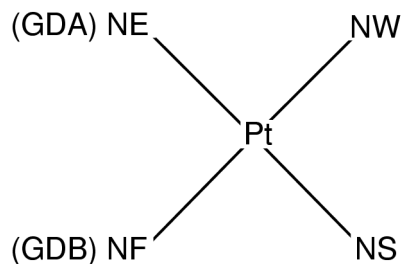


FIG. 2.1 – Modèle de représentation du cisplatine dans le champ de force d'AMBER.

TAB. 2.1 – Caractéristiques majeures de la paramétrisation du cisplatine dans le champ de force d'AMBER. HT : atome d'hydrogène relié aux ammines platinées. Le détail de chaque résidu est présenté en annexe, page 257.

Atome/résidu	R^* (Å)	ϵ / charge globale (électron)
Pt	2,30	0,047440
HT	1,10	0,367553
NS/NW	1,82	-0,701331
GDA/GDB		-0,425045

Liaison	(Å)	k (kcal.mol ⁻¹ .Å ⁻²)
Pt-NE	2,00	202,00
Pt-NF	2,00	202,00
Pt-NS	2,04	202,00
Pt-NW	2,04	202,00
NS/W-H	1,01	434,00

Angle	(°)	k (kcal.mol ⁻¹ .Å ⁻²)
NE - Pt - NS	180,0	20,00
NF - Pt - NW	180,0	20,00
NE - Pt - NW	90,0	40,00
NF - Pt - NS	90,0	40,00
NE - Pt - NF	90,0	40,00
NS - Pt - NW	90,0	40,00

Angle impropre	(°)	k (kcal.mol ⁻¹)
PT-C5-C2-N7	180,0	2,50
PT-C5-C2-C8	180,0	2,50
PT-C8-C5-C4	180,0	15,90
PT-C8-C6-C5	180,0	14,98

Les aspects les plus délicats de la représentation dans le champ de force concernent l'interaction entre le cisplatine et les bases platinées, essentiellement sous la forme de la répartition des charges [106] et de l'angle formé entre la base chélatée et la liaison Pt-N₇. Ce dernier paramètre a été mis en place par l'utilisation d'angles « impropres » appelés ainsi car ils ne concernent pas 4 atomes reliés par une liaison de covalence [38, 60, 108].

Une approche similaire a été utilisée par l'équipe de Marzilli et ses collaborateurs pour modéliser ce comportement [211].

Chapitre 3

Déformations induites sur le squelette

Dans le laboratoire, l'introduction de la méthode Particle Mesh Ewald (PME) pour le calcul des forces électrostatiques sous conditions périodiques dans les simulations de dynamique moléculaire a permis de donner une description dynamique de la conformation de l'ADN platiné. De nombreuses observations issues des études expérimentales [49, 56, 71, 87, 129, 142, 181, 182, 183, 210], tout particulièrement celles provenant des expériences RMN, ont ainsi été expliquées.

En effet, le déblindage du proton H2' de la base située en 5' de GDA et celui plus faible du proton H2' de la base complémentaire de GDA, ne pouvaient être expliqués simultanément par les modèles statiques proposés précédemment. L'apport de la simulation de dynamique moléculaire a été dans ce domaine très important car elle a pu démontrer la dynamique du mouvement, qui donnait enfin une explication claire du phénomène observé. De plus, le passage du champ de force parm 94 [47] à parm 98 [29] a permis d'élucider complètement la structure et la dynamique de l'adduit platiné.

Ces simulations ont de plus apporté des informations supplémentaires. C'est ainsi que des transitions BI \rightleftharpoons BII des groupes phosphates en 3' de la guanine platinée ont été observées. Ces transitions sont reliées à la formation/rupture d'une liaison hydrogène entre un atome d'hydrogène de l'ammine 3' du cisplatine (dont l'azote est noté NS sur la figure 2.1) et l'atome N7 de l'adénine située en 3' du pontage cisPt-GG. Ces transitions ont de plus pu être proposées comme permettant une reconnaissance différentielle de cette sous-conformation par les protéines de type HMG [59].

Cependant lors d'analyses complémentaires, après publication de ces hypothèses sur les sous-conformations, des problèmes sont survenus. En prolongeant la simulation ainsi qu'en réalisant des simulations de séquences différentes, une des caractéristiques majeures des complexes de cisplatine, observée dans les études expérimentales, n'était plus remplie. Selon les études expérimentales, le sucre de la guanine 5' du pontage cisPt-GG est toujours dans la conformation N tandis que celui des différentes dynamiques présentait des transitions entre les états N et S. Il est à noter que ce problème n'était

jamais survenu dans les simulations utilisant parm 94. Il a donc fallu déterminer la cause de ce problème dans la paramétrisation.

La paramétrisation utilisée jusqu'alors ne concernait que les influences locales que l'adduit induit sur les bases, sans prendre en compte exclusivement le sucre et le squelette phosphodiester. Or le changement majeur qui nous concerne entre le champ de force parm 94 et parm 98 concerne la valeur de la constante de torsion de l'angle O-C-C-O qui détermine la barrière de transition entre les deux états N et S [201].

De plus, le champ de force parm 94 était connu pour déplacer l'équilibre vers la conformation N [29], ce qui favorisait la conformation observée expérimentalement pour la guanine 5' du chélate cisPt-GG. Le passage à parm 98 dans la version 6.0 d'AMBER a été effectué afin de mieux rendre compte de la dynamique des acides nucléiques mais a enlevé cet avantage énergétique en faveur de la conformation N.

Jusqu'alors l'effet électronique dû à la coordination était supposé limité strictement à la guanine dans la paramétrisation. Cependant à la suite de nos constatations sur les dynamiques moléculaires et de la publication de travaux réalisés par Plavec et ses collaborateurs [151] concernant la chélation d'un adduit platiné sur une guanine (2'-deoxy/riboxy-guanosine 3',5' bis(éthyl phosphate)), il a été démontré que l'effet anomère dû au cisplatine se répercute jusqu'au phosphate du nucléotide, induisant en particulier une stabilisation de la conformation *trans* de l'angle ϵ . Une telle stabilisation va de pair avec une proportion plus importante de la conformation BI.

Pour représenter au mieux les caractéristiques d'un complexe ADN-platine, la paramétrisation du sucre de la guanine platinée et d'une partie du phosphodiester adjacent a été modifiée en faisant varier la constante de torsion de l'angle O-C-C-O pour le sucre et celle de l'angle C-C-O-P pour epsilon, indépendamment et en conjonction.

3.1 Présentation des systèmes d'étude

Le complexe pyrazolate biplatiné présenté dans le chapitre IV a été choisi pour la mise au point des nouveaux paramètres. La déformation structurale induite par ce composé platiné est modérée, contrairement au cisplatine les effets électroniques induits par la platination des guanines ne sont plus masqués par la coude. De plus, celui-ci comporte deux atomes de platine, un par guanine platinée, ce qui permet de vérifier indépendamment l'influence de la paramétrisation. Dans le cadre de ce complexe, les déformations structurales induites sur les guanines platinées sont différentes de celles du cisplatine : les informations issues de l'expérimentation RMN indiquent que le sucre de la guanine 3' est toujours en conformation N tandis que celui de la guanine 5' peut adopter les deux conformations N et S.

Pour l'étude de l'angle ϵ une simulation dérivée du travail de Elizondo-Riojas et

Kozelka [59], présentée dans le chapitre II, a été utilisée. Il s'agit de la séquence « mutée » par rapport au travail publié, la base suivant les deux guanines platinées étant une guanine au lieu d'une adénine. Cette « mutation » a été effectuée pour valider l'hypothèse émise par Elizondo-Riojas et ses collaborateurs concernant l'existence d'une liaison hydrogène entre NS et l'atome N_7 de l'adénine en 3', cette liaison étant couplée à l'apparition de la conformation BII. Cette conformation est très sensible à la conformation de l'angle ϵ et permet donc d'étudier l'influence de notre paramétrisation sur l'équilibre $BI \rightleftharpoons BII$.

3.2 Le sucre platiné

Pour distinguer les sucres d'une base platinée de ceux des bases non modifiées, il faut définir un nouveau type d'atome pour le cycle furanose. La figure 3.1 présente la localisation des changements nécessaires : le type d'atome CT, devient CP (P pour platiné). La modification n'est effectuée que pour les carbones constitutifs du cycle ($C_{1'}$, $C_{2'}$, $C_{3'}$, $C_{4'}$), les paramètres des atomes (longueur de liaison, angle, angles de torsion, charges) sont conservés identiques à ceux de parm 98 à l'exception de l'angle de torsion unique **OS-CP-CP-OS**. L'équation 3.1 présente la forme générique du potentiel pour un angle de torsion tandis que l'équation 3.2 présente son implémentation dans le calcul du champ de force.

$$E_{torsion} = \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \quad (3.1)$$

$$E_{torsion} = \frac{PK}{IDIVF} [1 + \cos(PN * \phi - PHASE)] \quad (3.2)$$

Pour la définition de l'angle de torsion **OS-CP-CP-OS** les valeurs sont respectivement¹ :

A-B-C-D	IDIVF	PK	PHASE	PN
OS-CT-OS-CT	1	0,144	0,0	-3
OS-CT-OS-CT	1	1,50	0,0	2

Pour améliorer le comportement du sucre platiné la valeur de la constante PK_{OCCO} du terme de périodicité $PN=2$ a été modifiée. Il est dans le champ de force parm 98 égal à 1,50 kcal.mol⁻¹.

L'étude des modifications du potentiel en faisant varier la constante de torsion PK_{OCCO} a été réalisée sur le complexe pyrazolato bis-platine présenté ci-dessus et dans le chapitre IV.

¹Les deux potentiels sont pris en compte et additionnés grâce à l'utilisation du signe négatif pour PN (-3 au lieu de 3).

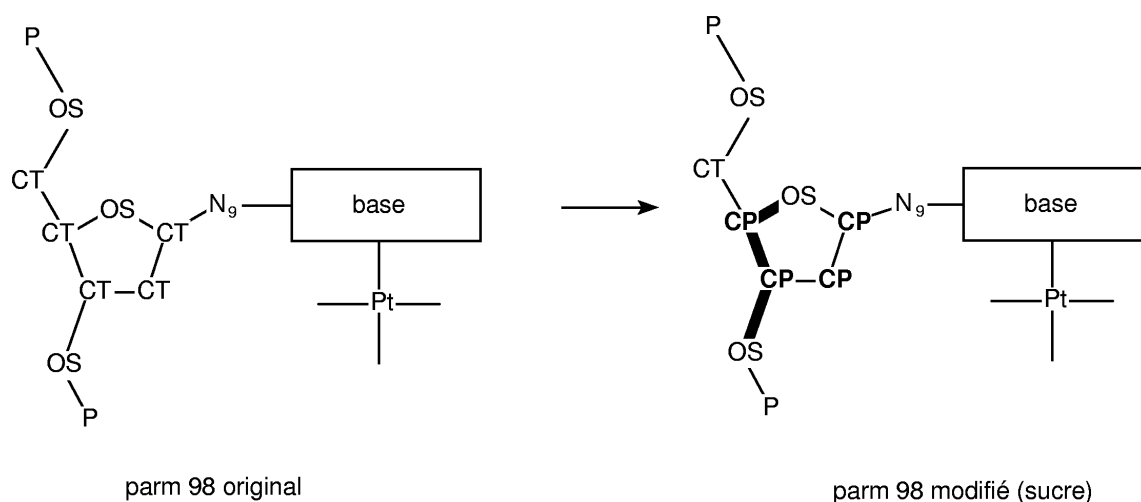


FIG. 3.1 – Modification apportée pour singulariser le sucre de la guanine platinée. Le type d'atome CT devient CP pour les atomes du cycle.

Dans un premier temps, la valeur de $0,75 \text{ kcal.mol}^{-1}$ a été retenue car elle correspond à la valeur de parm 94 (1.0) moins la moitié de la différence entre parm 94 et parm 98 (1.50). C'est une valeur qui permet de favoriser la conformation N des sucres [29].

Comme présenté dans le chapitre IV, trois critères issus des observations RMN sont discriminants pour le choix de la paramétrisation de la constante PK_{OCCO} :

1. la phase de la guanine platinée 3' doit être très majoritairement en N ;
2. la phase de la guanine platinée 5' est majoritairement S ;
3. la distance ${}^5\text{G-H}_1$ - ${}^3\text{G-H}_1$ doit être comprise entre 2,5 et 3,5 Å, la mesure nOe observée en RMN est qualifiée de « *medium* ».

Trois dynamiques ont été étudiées avec la constante pour l'angle de torsion O-C-C-O égale à 1,50 (parm 98), 0,75 (celle choisie arbitrairement) et 0,25 kcal.mol^{-1} (pour vérifier l'influence de la paramétrisation extrême). Les résultats sont présentés dans la figure 3.2.

Pour la constante $PK_{OCCO} = 0,75 \text{ kcal.mol}^{-1}$ (figure 3.2 a) : cette valeur ayant été choisie de manière arbitraire pour le point de départ de la simulation il advient de l'étudier avec plus de circonspection. Dans le cadre de la nanoseconde simulée, les trois pré-requis expérimentaux sont remplis : le sucre de la guanine ${}^5\text{G}$ adopte une conformation S, celui de la guanine ${}^3\text{G}$ une conformation N et la distance ${}^5\text{G}(\text{H}_1)$ - ${}^3\text{G}(\text{H}_1)$ est compatible avec une intensité RMN intermédiaire (« *medium* »). Cette valeur a donc été choisie pour effectuer une simulation plus longue, retenue pour produire un modèle à partir des indications RMN (comme présenté dans le chapitre IV). Il apparaît au terme de celle-ci que la guanine en 3' ne reste pas parfaitement en conformation N

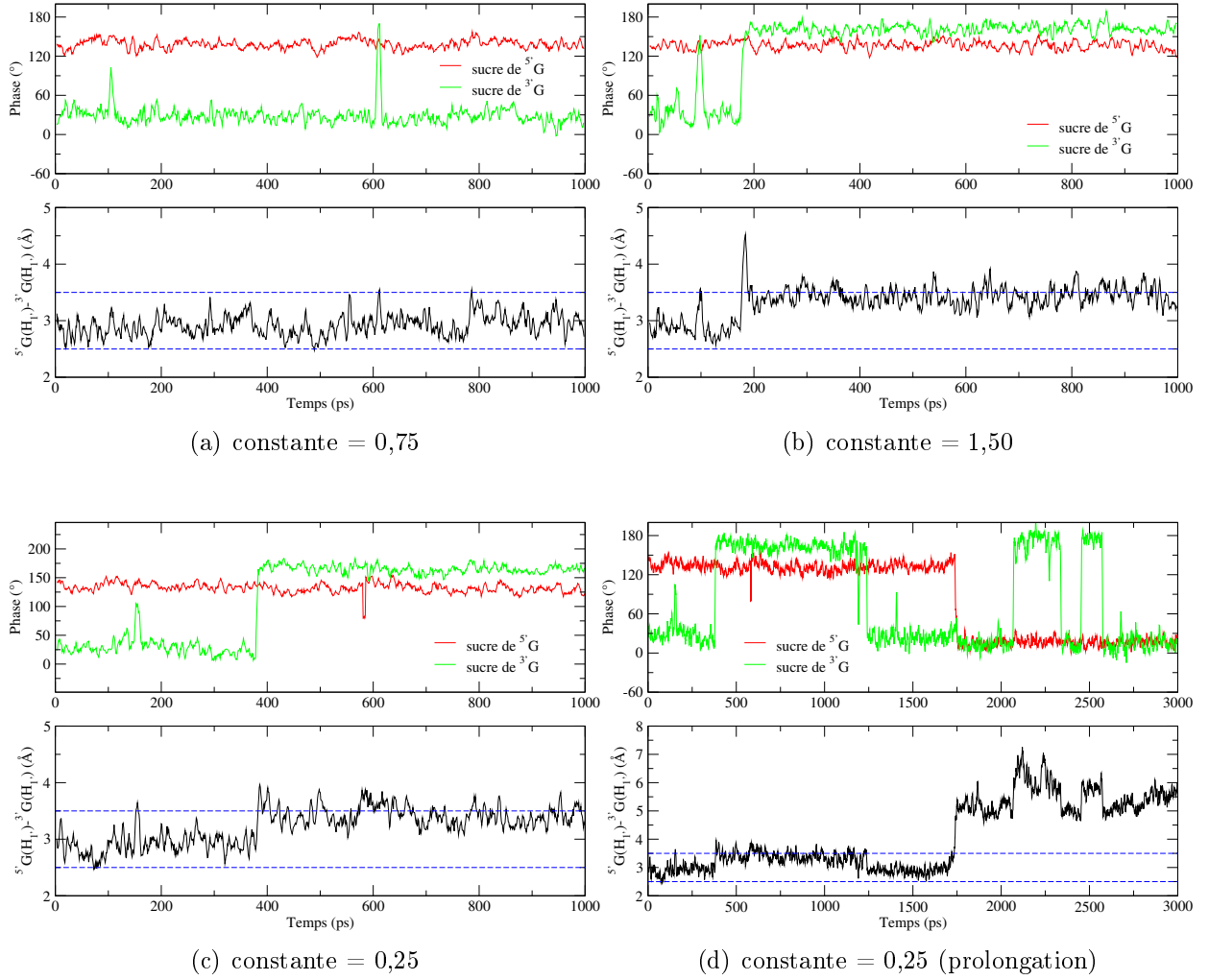


FIG. 3.2 – Influence de la valeur de la constante pour l'angle de torsion O-C-C-O sur la phase de $5'G$ (rouge), la phase de $3'G$ (vert) et la distance $5'G(H_1) - 3'G(H_1)$ (noir). Les tracés sont moyennés sur 5 ps pour plus de lisibilité.

et que la guanine 5' ne présente pas de transition $S \rightleftharpoons N$. Cela pourrait être l'indicateur qu'il faut encore diminuer la valeur de 0,75 pour se rapprocher du comportement observé avec la constante égale à 0,25 kcal.mol⁻¹ (voir ci-dessous). Cependant lors de l'analyse de la deuxième moitié de la simulation (10 ns), la proportion N/S pour la guanine 3' s'établit à 65 % N / 35 % S (elle est de 65% S / 35 % N pour les 10 premières nanosecondes). La valeur choisie pour la constante PK_{OCCO} pourrait donc être suffisante avec un point de départ de simulation proche du point d'équilibre, une simulation plus longue permettrait de le vérifier. Comme nous n'avons pas de données expérimentales suffisamment précises pour affiner ce paramètre, la constante de 0,75 kcal.mol⁻¹ est donc conservée pour modéliser l'effet anomère dû à la chélation.

Pour la constante $PK_{OCCO} = 1,50 \text{ kcal.mol}^{-1}$ (figure 3.2 b) (parm 98 natif) : après une brève incursion à 100 ps de la guanine 3' vers la conformation S, la transition devient définitive à 170 ps. La distance de référence reste encore une fois dans la marge haute acceptable mais par opposition à la simulation avec la constante égale à 0,25, les deux sucres restent en S, ce qui exclut définitivement cette valeur (voir ci-dessous).

Pour la constante $PK_{OCCO} = 0,25 \text{ kcal.mol}^{-1}$ (figure 3.2 c) : au début de la dynamique, les trois critères sont remplis. Cependant à 375 ps le sucre $3'G$ passe en conformation S. A ce moment la distance $5'G-H_1-3'G-H_1$ se retrouve dans la limite haute de la fourchette. Les deux sucres devraient tendre vers la conformation N puisque la paramétrisation incite cette conformation, or ceux-ci passent tous les deux en S. Après prolongation de cette dynamique (figure 3.2 d), il s'avère que la conformation N/N est bien adoptée mais après passage par des états intermédiaires (notamment avec ϵ et ζ en conformation très inhabituelle *gauche* -, *gauche* -). Cette transition apparaît autour de 1700 ps de simulation de dynamique et confirme que la paramétrisation permet de faire évoluer la conformation du sucre vers l'équilibre observé expérimentalement. Il est à noter que lorsque cette transition survient, la distance de référence devient très grande (supérieure à 5 Å) ce qui exclut définitivement cette conformation compte-tenu des paramètres expérimentaux.

A travers l'étude de trois valeurs différentes de la constante de torsion telle que définie plus haut, il apparaît que le champ de force parm 98 ne permet pas de rendre compte correctement des déformations induites par le pyrazolate bisplatine sur l'ADN. La modification effectuée permet par contre d'y arriver en utilisant la constante de 0,75 kcal.mol⁻¹ et une constante trop faible déplace l'équilibre $N \rightleftharpoons S$ de manière disproportionnée.

3.3 L'angle ϵ

Comme décrit par Plavec et ses collaborateurs [151], lors de la coordination d'une guanine (2'-deoxy/riboxy-guanosine 3',5' bis(éthyl phosphate)) au cisplatine l'effet anomère se répercute jusqu'au phosphate du nucléotide, induisant en particulier une stabilisation de la conformation *trans* de l'angle ϵ . Une telle stabilisation va de pair avec une proportion plus importante de la conformation BI. Pour prendre explicitement en compte cet effet, un nouveau type d'atome, CS a été ajouté uniquement pour le carbone $C_{2'}$, comme illustré sur la figure 3.3. L'angle de torsion ϵ est défini par les atomes $C_{4'}-C_{3'}-O_{3'}-P$ alors que nous utilisons ici $C_{2'}-C_{3'}-O_{3'}-P$ (décalé de 120° par rapport à ϵ , voir remarque plus bas), ce choix a été fait pour éviter d'interférer avec l'angle $C_{5'}-C_{4'}-C_{3'}-O_{3'}(\delta)$ qui a déjà été modifié par l'utilisation de la constante de torsion PK_{OCCO}

= 0,75 kcal.mol⁻¹ sur O-C-C-O.

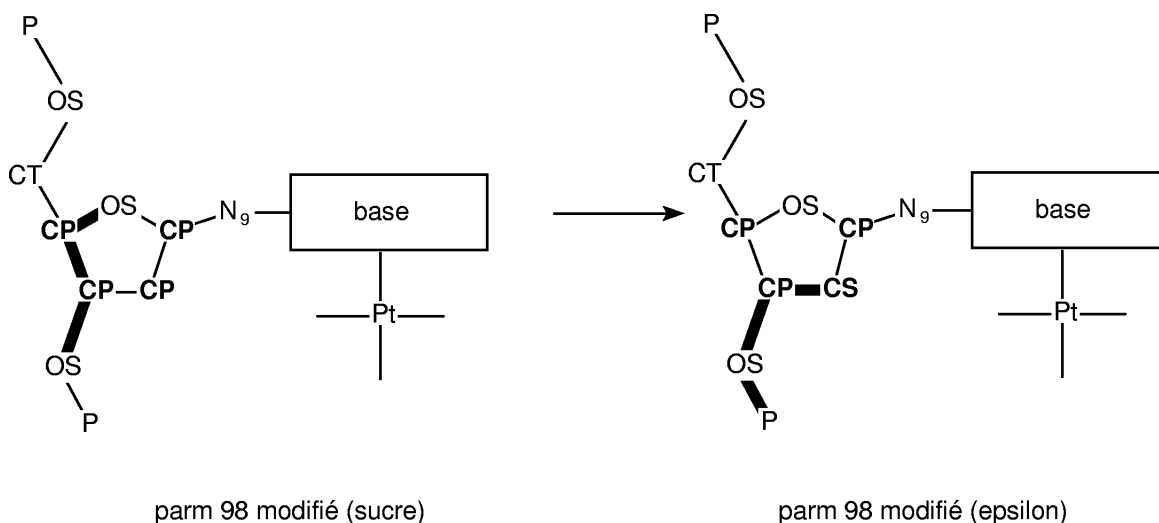


FIG. 3.3 – Modification apportée pour influencer uniquement sur l'angle ϵ . Le carbone C₂' est maintenant de type CS.

Dans ce cas il ne s'agit pas de modifier la valeur d'une constante existante (PK_{OCCO}) mais bel et bien d'en introduire un nouveau terme de torsion dans le champ de force. Ainsi les modifications se doivent d'être graduelles pour évaluer l'impact de ce nouveau terme de torsion.

De plus il faut favoriser l'état BI par rapport à l'état BII. Le moyen le plus efficace consiste à modifier le potentiel pour qu'au terme initial de torsion (PN=3) soit ajouté un nouveau terme (PN=2) ne pénalisant pas énergétiquement la conformation à favoriser. Ainsi en augmentant la constante de force $PK_{CCOP} = x$ de ce potentiel additionnel (décrit ci-dessous), il est possible d'augmenter la barrière de transition entre les états BI et BII mais aussi défavoriser l'état BII en lui-même en le pénalisant énergétiquement.

Pour la définition de l'angle de torsion **CP-CP-OS-P** les termes de l'équation 3.2 sont respectivement :

A-B-C-D	IDIVF	PK	PHASE	PN
X-CT-OS-X	3	1,15	0,0	3

Ce potentiel est modifié pour inclure un terme spécifique pour les sucres des nucléotides platinés. Les types d'atomes spécifiques sont alors CP et CS, et la barrière de torsion est définie par deux termes, le premier est la copie du potentiel général, le second a un décalage de phase de -60° afin d'obtenir le minimum pour la conformation BI. Le résultat de la modification du potentiel est présenté ci-dessous. La variable x permet d'adapter la constante du potentiel.

A-B-C-D	IDIVF	PK	PHASE	PN
CS-CP-OS-P	3	1,15	0,0	-3
CS-CP-OS-P	1	x	-60,0	2

Les modifications de potentiel qu'induit ce nouveau terme de torsion sont présentés dans la figure 3.4. Il est à noter que la conformation BI (qui se trouve ici à 60° , voir remarque plus bas) ne subit pas de pénalisation de l'énergie en fonction de $PK_{CCOP} = x$ tandis que c'est le cas pour la conformation BII (à 180° , voir remarque plus bas). La barrière de transition entre les deux états $BI \rightleftharpoons BII$ est elle aussi augmentée. Cette modification du potentiel stabilise la conformation BI, à la fois thermodynamiquement en défavorisant énergétiquement la conformation BII par rapport à la conformation BI et cinétiquement en augmentant la barrière de transition entre les deux états.

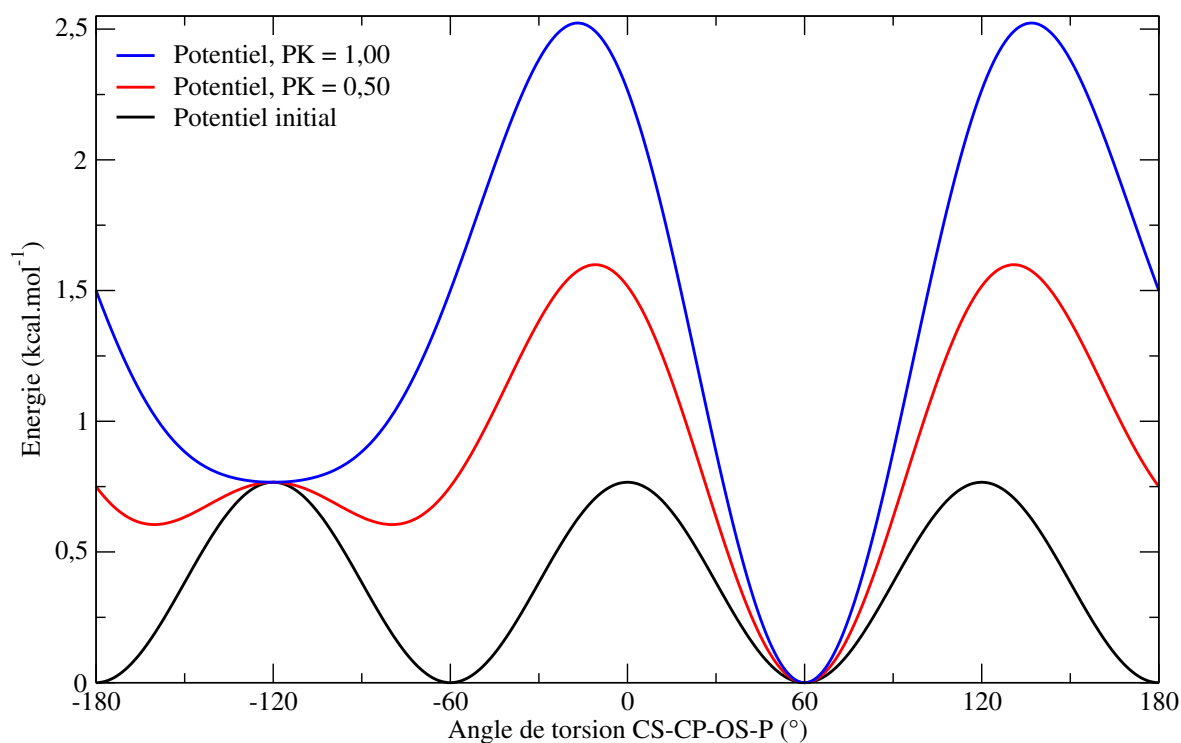


FIG. 3.4 – Evolution du potentiel d'énergie de torsion autour de l'angle CS-CP-OS-P en fonction de la constante PK_{CCOP} d'énergie utilisée.

Remarque : l'angle de la figure 3.4 est CS-CP-OS-P, pour obtenir la valeur de l'angle ϵ il faut ajouter 120° . Ainsi l'état le plus stable pour cet angle est $60+120 = 180^\circ$ (conformation *trans* présente dans l'état BI) suivi par l'état *gauche-* à $180+120 = 300^\circ$ que l'on cherche à défavoriser ici.

L'analyse par RMN de l'oligonucléotide platiné sur la séquence $5'-G^*G^*G-3'$ (cf. chapitre II) indique que le sucre de la guanine 5' adopte la conformation N (ce qui sous-tend une conformation BI) tandis que le phosphate de la guanine 3' est en conformation BI.

Pour reproduire ce comportement et étalonner la paramétrisation du nouveau terme, un point de départ identique issu de la simulation de dynamique moléculaire du cisplatine fixé sur la séquence centrale $5'-G^*G^*G-3'$ a été utilisé. Dans cette simulation la conformation initiale a comme caractéristiques un sucre S pour la guanine 5' et une conformation très majoritairement BII ($\epsilon_{3'G} - \zeta_{3'G} > 0^\circ$) pour le phosphate de la guanine 3'.

La constante de torsion $PK_{OCCO} = 0,75 \text{ kcal.mol}^{-1}$ sur O-C-C-O a tout d'abord été évaluée pour vérifier l'impact seul de cette modification sur la conformation de l'oligonucléotide platiné. Le nouveau terme a été étalonné avec les valeurs de $PK_{CCOP} = 0,50$ et $PK_{CCOP} = 1,00 \text{ kcal.mol}^{-1}$. L'évolution de la conformation en utilisant la modification sur ϵ ($PK_{CCOP} = 1,00 \text{ kcal.mol}^{-1}$) sans la modification sur les sucres platinés ($PK_{OCCO} = 1,50 \text{ kcal.mol}^{-1}$) a aussi été étudiée.

Evolution de la dynamique sans modification (figure 3.5 a) : les 200 premières ps délimitées par le trait vertical bleu représentent le point de départ issu d'une simulation utilisant parm 98 sans modification. Cette simulation a duré jusqu'alors 4 ns et les 200 ps présentées sont représentatives du comportement des deux guanines lors des précédentes 3,8 ns. La guanine $5'G$ est à ce moment de la dynamique dans la conformation S/BI (le sucre est en conformation S, le phosphate en BI ($\epsilon_{5'G} - \zeta_{5'G} < 0^\circ$)). Pour la guanine $3'G$ le sucre est aussi S, le phosphate présente des transitions BI \rightleftharpoons BII. Ces caractéristiques seront présentées systématiquement dans les graphiques suivants pour rappel. L'ensemble des résultats présentés ici (5*2 ns) fera l'objet d'une étude plus extensive (5*20 ns).

La prolongation de cette dynamique pour 2 ns permet l'observation du comportement des guanines si aucune re-paramétrisation n'était intervenue. Le phosphate de la guanine 5' reste en permanence en BI tandis que le sucre présente une transition de S vers N après 1280 ps de prolongation de la dynamique, cette transition faisant l'objet d'une augmentation des valeurs de $\epsilon_{5'G}$ et $\zeta_{5'G}$. Cette transition S \rightleftharpoons N est réversible, puisque dans le prolongement de cette dynamique la conformation S réapparaîtra (les détails des 5*20ns feront l'objet d'une publication ultérieure). La conformation N qui est observée expérimentalement peut donc être observée en utilisant le champ de force parm 98 mais cet état n'est pas stable dans la durée de la simulation.

En ce qui concerne la guanine 3' le sucre reste dans la conformation S. Le phosphate présente des transitions BI \rightleftharpoons BII avec une proportion importante de BII.

L'utilisation du champ de force parm 98 permet de s'approcher des observations expérimentales (issues d'études antérieures et de celle présentée dans le chapitre II) mais les états caractéristiques ne sont pas présents car ils oscillent entre deux états métastables.

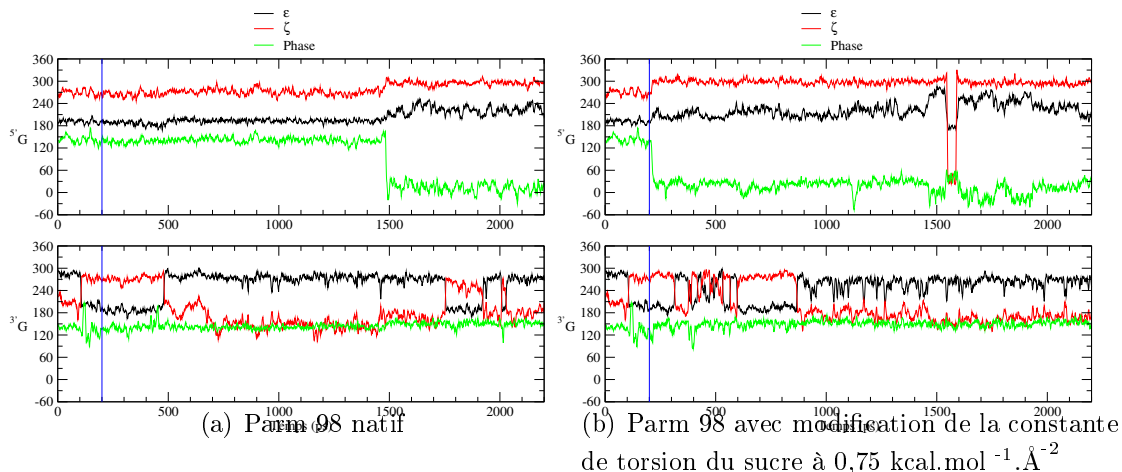


FIG. 3.5 – Prolongation de la simulation de dynamique moléculaire avec parm 98. La phase des sucres est représentée en vert, l'angle de torsion ϵ est en noir, l'angle de torsion ζ est en rouge. Les 200 premières ps délimitées par la ligne verticale bleue présentent le comportement dynamique avec parm 98 natif. Les tracés sont moyennés sur 5 ps pour plus de lisibilité.

Evolution en modifiant uniquement le potentiel de torsion des sucres platinés ($PK_{OCCO} = 0,75 \text{ kcal.mol}^{-1}$, figure 3.5 b) : en ajoutant au champ de force la modification des sucres platinés mise au point plus haut, le comportement du sucre de la guanine 5' évolue. Au lieu de présenter des transitions $S \rightleftharpoons N$ la conformation devient rapidement N et le reste tout au long de la simulation. Il est à noter que la conformation N va de pair avec des valeurs pour $\epsilon_{5'G}$ et $\zeta_{5'G}$ plus élevées que lorsque l'on a la conformation S. De plus cette modification n'influe pas sur la phase du sucre de la guanine 3' qui reste S.

Le phosphate de la guanine 5' est en conformation BI, comme précédemment le phosphate de la guanine 3' est majoritairement BII.

La modification de la constante de torsion des deux sucres platinés permet donc sélectivement d'approcher la conformation du sucre de la guanine 5' observée expérimentalement sans altérer celle du sucre de la guanine 3'. Cette reparamétrisation n'influe pas sur le comportement du phosphate.

Evolution en modifiant le potentiel de torsion des sucres platinés et le potentiel de torsion du terme pour ϵ ($PK_{OCCO} = 0,75 \text{ kcal.mol}^{-1}$, $PK_{CCOP} = 1,00 \text{ kcal.mol}^{-1}$, figure 3.6) : comme indiqué plus haut, l'angle $C_{2'}-C_{3'}-O_{3'}-P$ a été préféré à l'angle $C_{4'}-C_{3'}-O_{3'}-P$ pour éviter l'utilisation d'atomes identiques pour deux potentiels différents. Ces deux potentiels sont en partie couplés comme l'indique la période de latence pour la transition de l'état S vers N pour le sucre de la guanine

5' : 250 ps après la prolongation sont nécessaires pour observer cette transition avec $PK_{OCCO} = 0,75$ kcal.mol⁻¹, $PK_{CCOP} = 1,00$ kcal.mol⁻¹ tandis qu'elle est immédiate avec l'utilisation seule de la constante $PK_{OCCO} = 0,75$ kcal.mol⁻¹ pour le potentiel de torsion de ce sucre (figure 3.5 b). Cette période de latence écoulee, la transition vers N est permanente.

Pour la guanine 3' le sucre est toujours S, le phosphate adopte maintenant la conformation BI (avec une courte transition vers BII entre 390 et 450 ps après prolongation).

Cette nouvelle paramétrisation utilisant une combinaison de la modification des sucres et de ϵ permet d'obtenir un ensemble dynamique de structures reproduisant les caractéristiques essentielles observées en RMN et radiocristallographie aux rayons X.

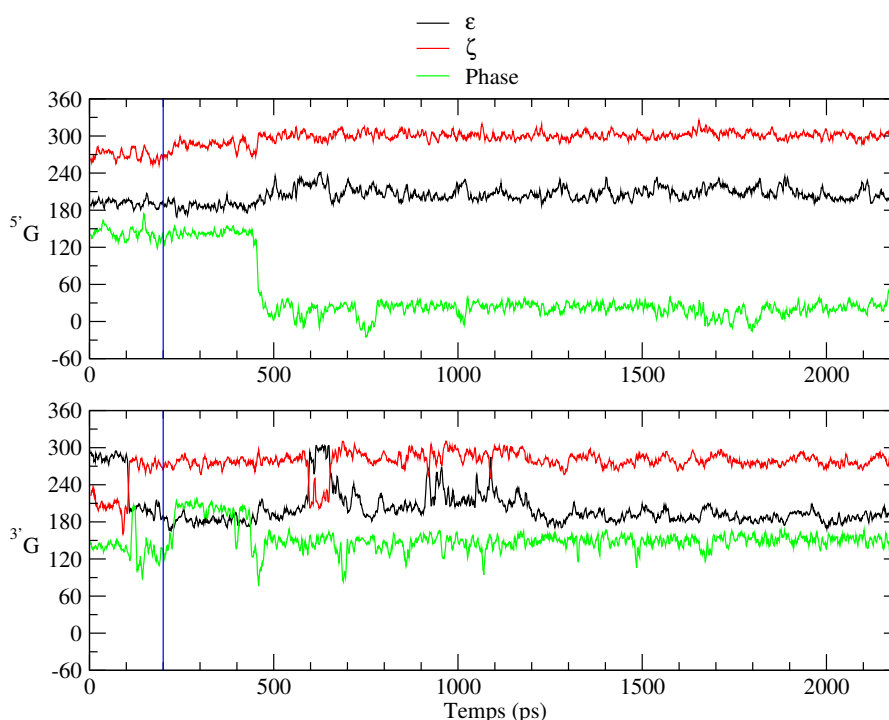


FIG. 3.6 – Prolongation de la simulation de dynamique moléculaire avec parm 98 puis modification des sucres platinés et du potentiel de torsion avec $PK_{CCOP} = 1,00$ kcal.mol⁻¹ pour ϵ . La phase des sucres est représentée en vert, l'angle de torsion ϵ est en noir, l'angle de torsion ζ est en rouge. Les 200 premières ps délimitées par la ligne verticale bleue présentent le comportement dynamique avec parm 98 natif. Les tracés sont moyennés sur 5 ps pour plus de lisibilité.

Evolution en modifiant le potentiel de torsion des sucres platinés et le potentiel de torsion du terme pour ϵ ($PK_{OCCO} = 0,75$ kcal.mol⁻¹, $PK_{CCOP} = 0,50$ kcal.mol⁻¹, figure 3.7) : avec le terme pour ϵ de moitié inférieur à celui illustré par la figure 3.6, la période de latence est très réduite pour le passage du sucre de S vers

N (25 ps au lieu de 250 ps). La conformation du sucre et du phosphate de la guanine 3' sont compatibles avec les observations expérimentales.

Cette paramétrisation pourrait être suffisante pour rendre compte des effets électroniques induits pas la chélation, cependant la conformation du phosphate de la guanine 5' est inhabituelle. En effet, l'angle ζ n'adopte pas une conformation canonique (g^- ou t) mais $gauche^+$. Après observation attentive des structures présentant cette caractéristique particulière il apparaît que la conformation globale obtenue est celle proposée précédemment au laboratoire : modèle E de Kozelka et ses collaborateurs [87, 107] où la cytosine complémentaire de la guanine 5' G reste empilée sur la cytosine complémentaire de la guanine 3' G plutôt que d'être appariée à celle-ci ou d'osciller entre les deux positions comme proposé par Elizondo-Riojas et ses collaborateurs [59]. Cette simulation de dynamique moléculaire en eau explicite est la première à démontrer la possibilité de l'empilement de type E'' calculés par mécanique moléculaire et rouvre la discussion concernant les équilibres conformationnels induits par les pontages G^*G^* -cisplatine.

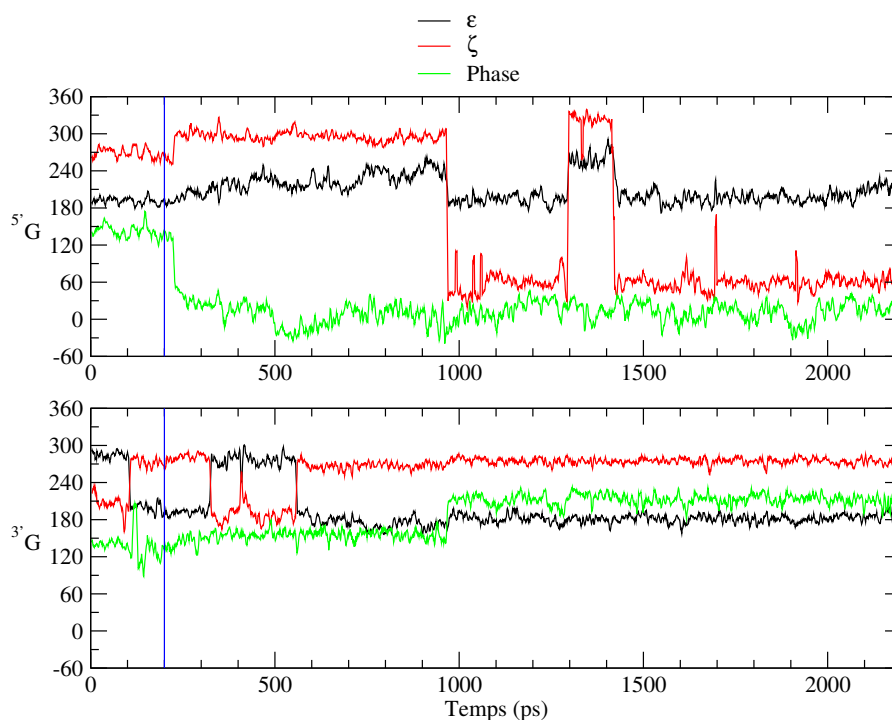


FIG. 3.7 – Prolongation de la simulation de dynamique moléculaire avec parm 98 puis modification des sucres platinés et du potentiel de torsion avec $PK_{CCOP} = 0,50$ kcal.mol $^{-1}$ pour ϵ . La phase des sucres est représentée en vert, l'angle de torsion ϵ est en noir, l'angle de torsion ζ est en rouge. Les 200 premières ps délimitées par la ligne verticale bleue présentent le comportement dynamique avec parm 98 natif. Les tracés sont moyennés sur 5 ps pour plus de lisibilité.

Evolution sans modification des sucres platinés, modification du potentiel de torsion du terme pour ϵ ($PK_{CCOP} = 1,00 \text{ kcal.mol}^{-1}$, figure 3.8) : comme vu précédemment la conjonction de la modification du sucre et l'utilisation de $1,00 \text{ kcal.mol}^{-1}$ pour ϵ conduit à une conformation répondant aux observations expérimentales. Pour étudier l'effet isolé de la modification du potentiel du phosphate (via ϵ) une simulation ne comprenant pas de modification sur les sucres platinés a été effectuée. Dans ces conditions le comportement général de la guanine 5' est identique au comportement observé pour la simulation avec parm 98 natif (la transition S vers N apparue nécessitera l'analyse des 20 ns pour déterminer plus précisément sa pertinence statistique). Une favorisation de la conformation BI est observable pour le phosphate de la guanine 3' indiquant bien un effet asymétrique de la paramétrisation identique des deux guanines. L'utilisation du nouveau terme pour la paramétrisation du phosphate ne permet donc pas à lui seul à faire converger les structures dynamiques vers les observations expérimentales.

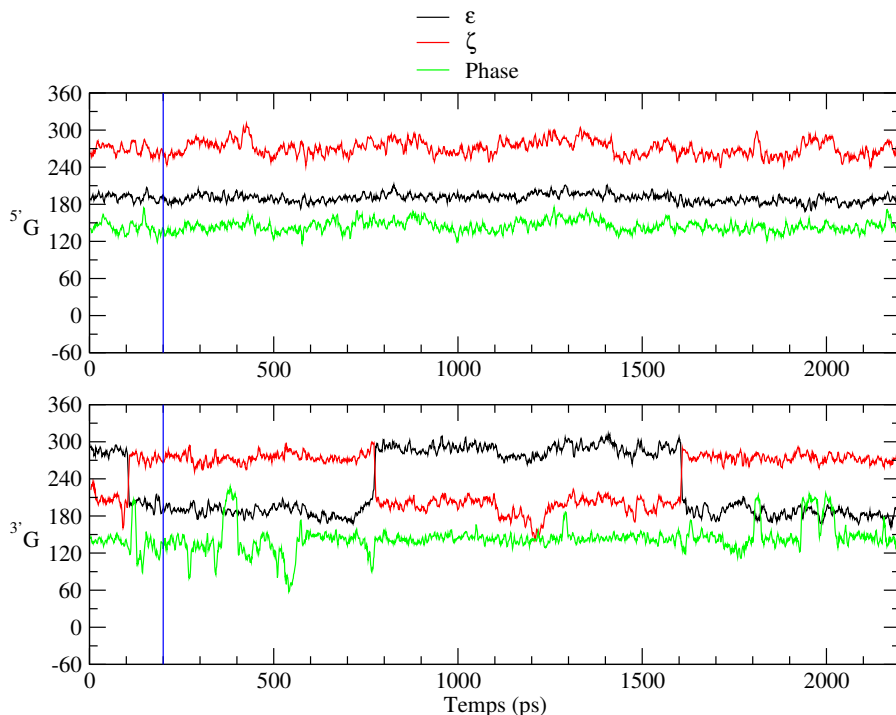


FIG. 3.8 – Prolongation de la simulation de dynamique moléculaire sous parm 98 avec le potentiel de torsion modifié $PK_{CCOP} = 1,00 \text{ kcal.mol}^{-1}$ pour ϵ mais sans modification des sucres platinés. La phase des sucres est représentée en vert, l'angle de torsion ϵ est en noir, l'angle de torsion ζ est en rouge. Les 200 premières ps délimitées par la ligne verticale bleue présentent le comportement dynamique avec parm 98 natif. Les tracés sont moyennés sur 5 ps pour plus de lisibilité.

La modification du champ de force parm 98 a donc nécessité, en plus de la paramétrisation déjà réalisée précédemment au laboratoire, l'inclusion explicite des effets

électroniques sur le sucre et le squelette phosphodiester. Les modifications présentées feront l'objet d'études plus approfondies pour tester leur validité sur d'autres séquences et complexes, mais sur ces deux cas complémentaires (présentés chapitres II et IV), ils ont permis d'obtenir un meilleur accord entre le modèle théorique et les expériences pratiques.

Chapitre 4

Intégration dans le champ de force

La présence d'angles impropres dans la paramétrisation du pontage G*G*-cisplatine prévient l'utilisation des modules évolués d'AMBER (xleap ou antechamber). Voici donc en détail la procédure à suivre, utilisée dès la création du logiciel AMBER [200], pour l'incorporation d'un composé platiné dans le champ de force. Le détail des paramètres de chaque fichier est présenté dans l'annexe, page 257. A titre d'exemple, voici la procédure permettant la mise en place des fichiers pour la simulation de dynamique moléculaire. Il s'agit de la séquence 5'-GCCG*G*GTCGC-3' / 5'-GCGACCCGGC-3' (G* représente une guanine platinée) avec le champ de force parm 99 [196]. Cette dynamique, comme présenté dans le chapitre II, a eu pour but d'utiliser une voie de paramétrisation alternative à celle présentée ci-dessus.

Les modifications ajoutées au champ de force pour prendre en compte la nouvelle paramétrisation décrite dans ce chapitre sont présentées en annexe, page 259.

4.1 Préparation du composé

Fichier prepin (page 261) : contient la description des différents résidus à insérer : GDA (guanine 5'), GDB (guanine 3') et PTJ (cisplatine).

Fichier linkin (page 264), linkin_ions (page 264) : les résidus sont reliés entre eux, la conformation est pour l'instant sous la forme de coordonnées internes.

Fichiers editin (page 264), editin_ions (page 264) et editin_eau (page 265) : la structure tridimensionnelle est affectée aux résidus, à partir d'un jeu de coordonnées externe (code PDB : 1AIO).

Fichier parmin (page 266) : les contraintes additionnelles sur les atomes sont ajoutées, dans notre cas, il s'agit de deux fois quatre angles impropres pour modéliser

l'influence de la chélation par le cisplatine sur les déformations des guanines.

Grâce aux fichiers décrits ci-dessus (prepin, linkin, editin et parmin) et à la modification du champ de force, la molécule est construite en exécutant la séquence PREP->LINK->EDIT->PARM (détail des scripts page 266) trois fois de suite, la différences majeures se jouant sur le fichier editin.

La première série de commandes se fait avec le fichier editin qui permet d'inclure le fichier de coordonnées tridimensionnelles, puis editin_ions est utilisé pour ajouter les contre-ions. Les molécules d'eau sont ajoutées par l'utilisation du fichier editin_eau, la distance minimale de la couche d'eau (entre le soluté et le bord de la boîte de solvation) étant ajustée pour que les dimensions de la boîte d'immersion ainsi créée soient compatibles avec l'utilisation de la Transformée Rapide de Fourier (la boîte doit avoir des dimensions multiples de 2, 3 ou 5 comme par exemple 64*54*60).

A ce stade la répartition des molécules d'eau forme une géométrie régulière, il convient donc d'optimiser leur position par rapport au soluté et aux contre-ions sodium. Cette opération est réalisée à travers une série de 6 minimisations où les contraintes de position sur l'ensemble soluté+contre-ions sont enlevées par paliers (100 kcal.mol⁻¹, 50, 25, 10, 5 sur le soluté seul, aucune). Les fichiers utilisés pour la minimisation sont présentés page 269.

4.2 Simulation de dynamique moléculaire

Le protocole décrit dans le tableau 4.1¹ détaille les étapes utilisées classiquement au laboratoire pour la simulation de la dynamique moléculaire d'un ADN platiné.

Il permet d'amener un ADN platiné de dix paires de bases en 67,503 ps à une conformation de départ pour la simulation de dynamique moléculaire.

Un script global permet de lancer toutes ces étapes à la suite (page 272).

4.3 Analyse de la dynamique

Des vérifications périodiques doivent être effectuées pour s'assurer du bon déroulement du calcul, et de la justesse de la simulation. Cependant, la puissance des ordinateurs est devenue telle qu'il s'agit souvent de vérifier plusieurs nanosecondes d'une journée sur l'autre. Dans ces conditions, il convient d'utiliser des outils d'analyse systématique pour avoir une représentation globale des paramètres importants. La figure 4.1

¹L'équilibration 9 est strictement identique à l'étape 8, mais elle a été ajoutée à l'origine pour permettre l'utilisation d'une version modifiée du programme AMBER 4.1 (corrigée par le docteur Matthew Young, afin de prendre réellement en compte la réinitialisation de la position du centre de masse). Elle est conservée pour permettre d'obtenir un point de départ ayant reçu le même traitement, mais les nouvelles versions (à partir d'AMBER 6.0) ne nécessitent plus ce point particulier.

TAB. 4.1 – Protocole de simulation de dynamique moléculaire.

Etape	Pas	Temps (ps)	Conditions thermodynamiques	Type de couplage au bain thermostaté	Contraintes (kcal.mol ⁻¹) ADN-Pt / Ions
Chauffage	1	5,0	NVT, 0K à 150K	facteur d'échelle	25 / 25
	2	5,0	NVT 150K à 300K	Berendsen	25 / 25
Equilibration	1	5,0	NVT, 300K	Berendsen	25 / 25
	2	2,5	NPT, 300K	Berendsen	25 / 25
	3	2,5	NPT, 300K	Berendsen	20 / 15
	4	2,5	NPT, 300K	Berendsen	15 / 10
	5	2,5	NPT, 300K	Berendsen	10 / 5
	6	2,5	NPT, 300K	Berendsen	5 / 2,5
	7	10,0	NPT, 300K	Berendsen	2,5 / -
Redist. vitesses	1	1 fs	300K	Berendsen	
Equilibration	8	10,0	NPT, 300K	Berendsen	
	9	10,0	NPT, 300K	Berendsen	
Redist. vitesses	2	1 fs	300K	Berendsen	
Equilibration	10	10,0	NPT, 300K	Berendsen	
Redist. vitesses	3	1fs	300K	Berendsen	
Production		x ns	NPT, 300K	Berendsen	

présente la décomposition de l'énergie du système simulé et d'autres points de contrôle (volume du système, marge d'erreur pour le calcul PME) pour le chauffage et l'équilibration de la dynamique. Les critères définis lors du lancement de la simulation sont conservés, que ce soit pour l'énergie totale ou la décomposition en sous-parties (potentielle et cinétique), après la phase initiale de chauffage. L'analyse de la dynamique complète avec la même méthodologie fournit le même résultat, mais les graphes sont moins lisibles du fait de la densité de points.

Dans le cas de l'ADN, il est possible d'utiliser le programme de référence CURVES [112, 113]. Afin d'automatiser la présentation graphique des données grâce au logiciel XmGrace (<http://plasma-gate.weizmann.ac.il/Grace/>), j'ai créé un programme d'analyse et de formattage, présenté en Annexe, page 288.

Enfin, il s'agit aussi de présenter les caractéristiques d'intérêt de l'adduit platiné (par exemple pour vérifier que la simulation est en accord avec les données expérimentales), comme illustré en partie par la figure 4.2 (pour plus de détails, cf chapitre II).

Ces vérifications étant faites, de nombreuses analyses particulières permettront de décrire au mieux l'accord de la simulation avec les données expérimentales et parfois de produire un modèle dynamique. C'est l'analyse et la réflexion qui seront alors

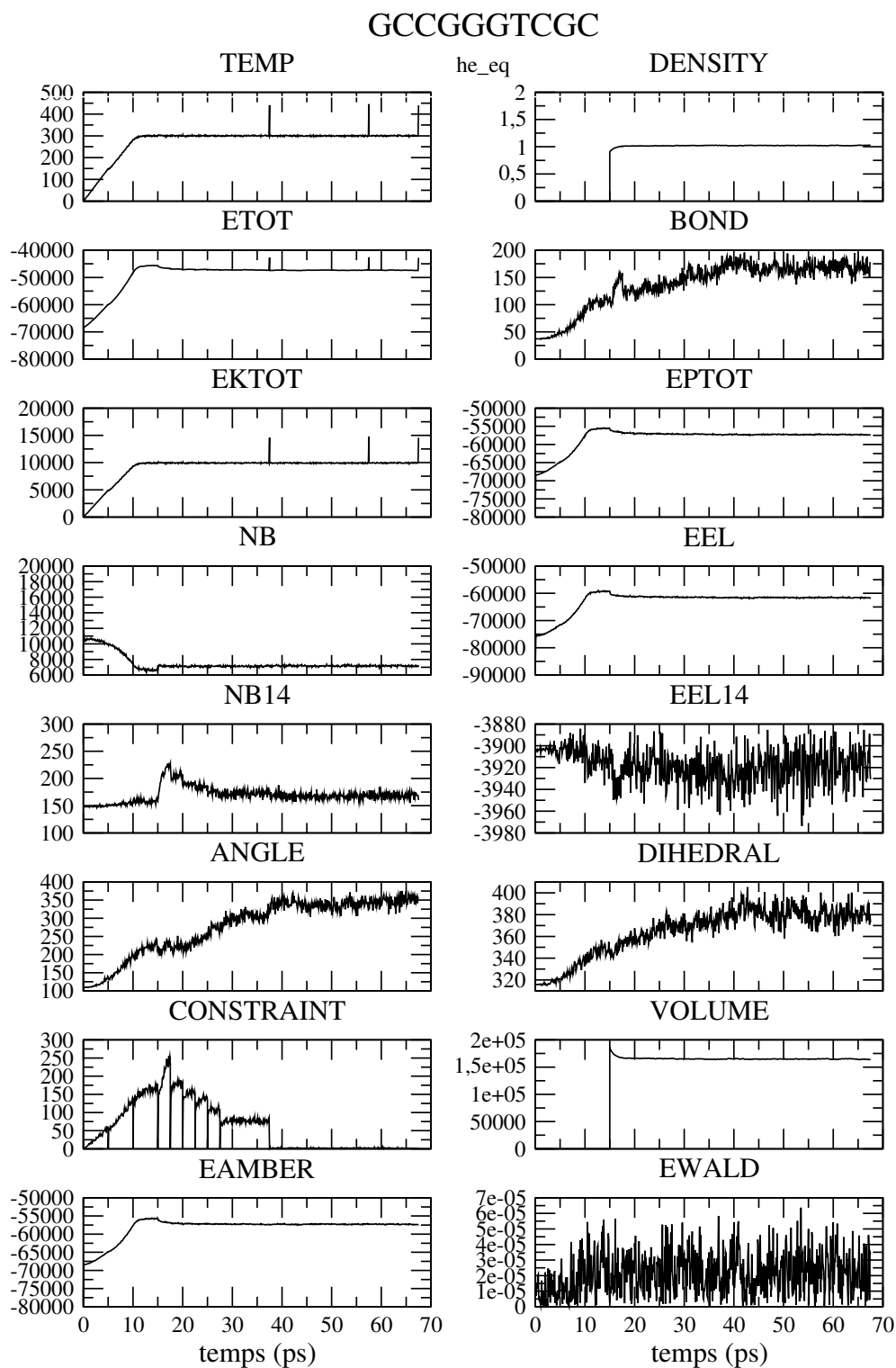


FIG. 4.1 – Contrôle du bon fonctionnement de la dynamique, en utilisant le programme standard `process_mdout.perl` fourni dans AMBER.

indispensables pour exprimer au mieux l'information présente dans la simulation.

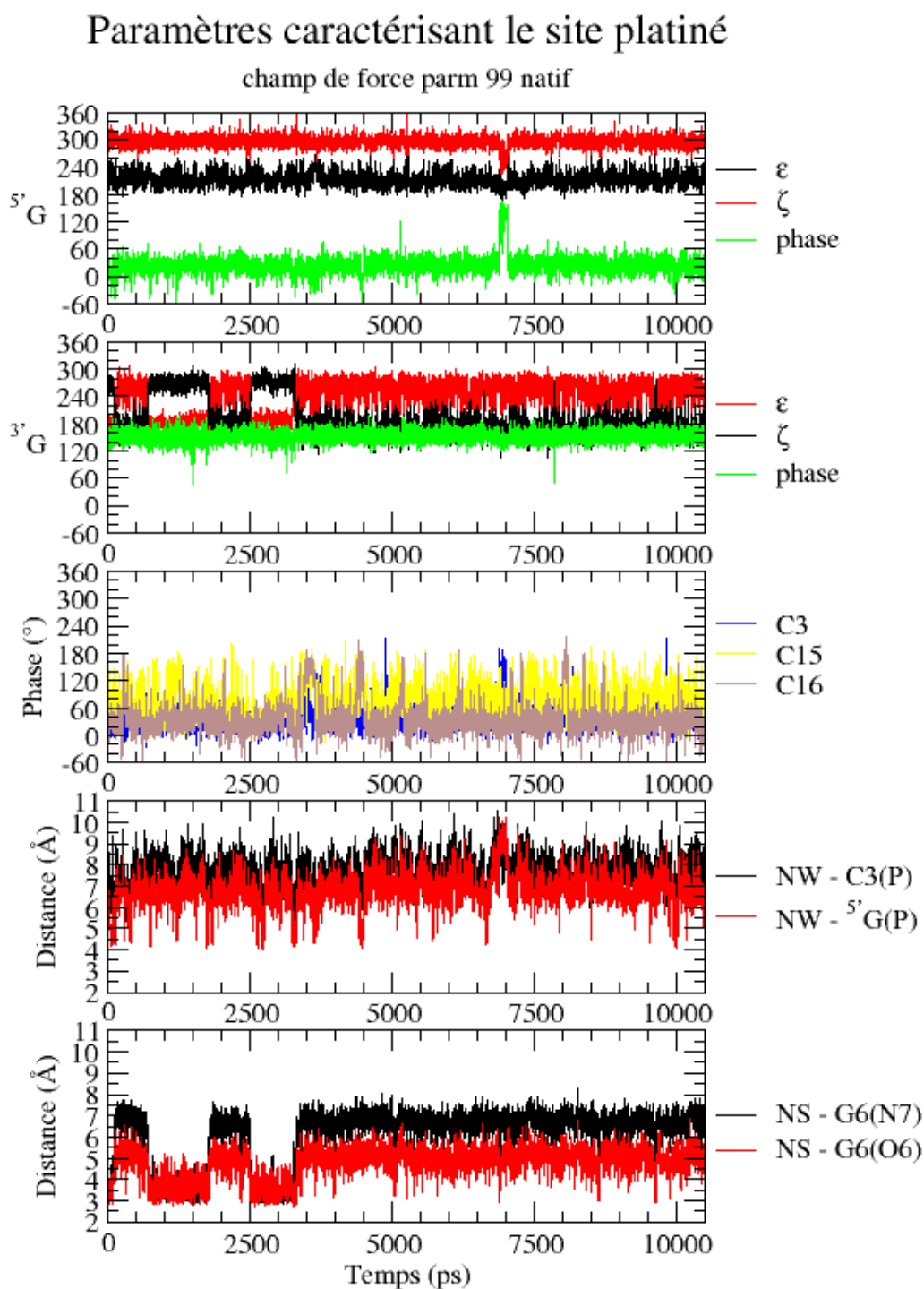


FIG. 4.2 – Paramètres principaux issus de la simulation de dynamique moléculaire de GCCG* G*GTCGC. Les noms d'atomes sont ceux de la figure 2.1, page 204. C3 et G6 sont respectivement les bases situées en 5' et en 3' de l'adduit platiné. C15 est la cytosine complémentaire de 3'G, C16 est la cytosine complémentaire de 5'G. Les comparaisons avec les données expérimentales sont présentées dans le chapitre II. Les tracés sont moyennés sur 10 ps pour plus de clarté.

Chapitre 5

Conclusion

Les composés de platine ne sont pas définis en standard dans le champ de force d'AMBER. Il faut donc décrire explicitement les interactions qui régissent leur comportement. De plus, la géométrie plan-carré du cisplatine impose l'utilisation d'angles impropres pour maintenir plan l'ensemble.

Les travaux effectués au laboratoire pour intégrer le cisplatine dans le champ de force ont été présentés. Cette paramétrisation a permis de répondre aux questions soulevées par l'analyse en spectroscopie RMN de l'oligonucléotide $5'GCCG*G*GTCGC^{3'}-5'GCGACCCGGC^{3'}$, tout particulièrement en ce qui concerne les déplacements chimiques H2' de C3 et C16.

Cependant la paramétrisation mise en place a volontairement exclu l'incorporation du sucre et du phosphate des guanines platinées. En effet l'influence des composés platinés sur ces résidus attachés aux guanines était supposée limitée. Les travaux de Plavec et de ses collaborateurs [151] ont indiqué que cette influence pouvait s'étendre jusqu'à l'oxygène O3' en induisant une stabilisation de la conformation N du sucre ainsi que de l'état *trans* pour ϵ , ce qui implique une conformation BI.

Des simulations de dynamique moléculaire en parallèle de ces travaux ont mis en évidence les faiblesses de la paramétrisation en ce qui concerne les sucres platinés et le potentiel de torsion pour ϵ . Deux complexes de platine induisant des déformations structurales différentes ont été utilisés pour la reparamétrisation. Le cisplatine courbe l'ADN et lui impose un angle de coude de 60° , le pyrazolato-bisplatine quand à lui courbe peu l'ADN mais impose une distance fixe entre les atomes N_7 des deux bases pontées. La complémentarité de ces déformations structurales a permis de valider par simulation de dynamique moléculaire les termes modifiés et/ou ajoutés pour rendre compte de l'effet de la coordination sur les sucres et le déplacement du phosphate.

Les résultats ainsi obtenus sont :

* pour le potentiel de torsion du sucre, l'angle OS-CP-CP-OS est altéré avec $PK_{OCCO} = 0,75 \text{ kcal.mol}^{-1}$ pour le sucre de la base platinée ;

* pour le potentiel de l'angle CS-CP-OS-P, le nouveau terme de torsion CP-CP-OS-P est créé avec pour valeur $PK_{CCOP} = 1,00 \text{ kcal.mol}^{-1}$ pour le squelette phosphodiester de la base platinée.

L'introduction de ces deux termes ainsi que leur étalonnage a permis de mieux rendre compte encore des données expérimentales, comme précisé dans le chapitre II. L'ensemble du nucléotide est maintenant caractérisé et la pertinence des modifications effectuées est évaluée dans le cadre de simulation plus longues (5*20 ns) qui feront l'objet d'une publication.

Sixième partie

Conclusion générale

Le cisplatine a depuis sa découverte été étudié en utilisant toutes les techniques disponibles en recherche expérimentale. Les tests chimiques, pharmacologiques, biochimiques et biologiques ont permis de découvrir les étapes clés de son efficacité thérapeutique. Les études structurales de RMN et de cristallographie ont servi à caractériser les structures formées au niveau moléculaire. Dans la continuité de ces travaux, nous nous sommes attachés à décrire le comportement dynamique de l'adduit ADN-cisplatine.

L'évolution informatique, matérielle et logicielle, permet en effet d'accéder à des temps de simulation de l'ordre de la dizaine de nanosecondes. Il devient possible d'utiliser la modélisation de dynamique moléculaire comme un outil pour caractériser les évolutions conformationnelles à petite échelle. Les mouvements observables durant cet intervalle de temps se cantonnent à des modifications locales, pour l'ADN il s'agit de variations dans le squelette (équilibres $BI \rightleftharpoons BII$, $N \rightleftharpoons S$) et entre les paires de base (désempilement, roulis, ...), pour les protéines, il s'agit des déplacements des chaînes latérales.

Observer expérimentalement des mouvements d'amplitude si faible est complexe : le cristal ne présente qu'une vue (ou deux) de la structure à l'état fondamental, la RMN manque de précisions sur les acides nucléiques compte-tenu des distances importantes qui séparent les protons et des phénomènes de diffusion de spin.

Pour améliorer les perspectives d'observation de la transition $BI \rightleftharpoons BII$, observée en dynamique moléculaire par le Dr. Elizondo-Riojas, nous avons mis au point une méthode plus robuste que les techniques actuelles pour l'analyse de la conformation du phosphate. En nous basant sur les distances caractéristiques des sous-populations BI et BII, obtenues à partir de modèles minimisés de dynamique moléculaire et de cristallographie, nous avons déterminé que quatre distances permettaient de discriminer précisément les deux sous-états. Pour fiabiliser la prédiction de la proportion de BII, nous avons combiné trois de ces distances en projetant la différence entre $H2''(n) - H6/H8(n)$ et $H2''(n) - H6/H8(n+1)$ contre $H1'(n) - H6/H8(n+1)$. Le graphe résultant de cette projection permet de caractériser un état BII avec une marge d'erreur de 20%. Une fois cette méthode validée par une série d'expériences RMN caractéristiques, elle pourra être incorporée dans les logiciels de raffinement RMN.

La conformation BII en 3' des guanines platinées ayant été observée en simulation de dynamique moléculaire, la validation expérimentale de cette conformation était nécessaire. J'ai donc platiné la séquence GCCGGGTCGC puis isolé et purifié les différents adduits. Les deux adduits majoritaires G^*G^*G et GG^*G^* étaient les premiers adduits étudiés comportant une guanine respectivement en 3' et en 5' du site de platination. Pour l'analyse en RMN, seul l'adduit G^*G^*G a pu être étudié, en raison de la trop faible concentration de l'adduit GG^*G^* . L'analyse RMN, effectuée en collaboration avec T. Skauge de l'équipe du Pr. Sletten, a permis de caractériser les sucres, les

phosphates et les distances intra et internucléotidiques. Ces résultats ont été couplés à une simulation de dynamique moléculaire mais la population BII observée par le Dr. Elizondo-Riojas n'a pas été retrouvée. Ce travail a cependant permis la détection de faiblesses de représentation des composés platinés dans le champ de force parm 98 de la suite AMBER. Depuis, ce champ de force a lui-même été corrigé pour améliorer les proportions des conformations N et S, en déplaçant l'équilibre vers N, ce qui est défavorable à la conformation BII.

L'étude expérimentale a de plus mis en évidence la migration du platine de l'adduit G*G*G vers l'adduit GG*G*. Une première approche de la cinétique de réarrangement a été présentée. L'étude cinétique spécifique effectuée par K. Chvalova est en cours et permettra de caractériser précisément les vitesses de réarrangement. Cette étude devra permettre de préciser le mécanisme de réarrangement. Toutefois les observations RMN suggèrent que la liaison se rompt entre le platine et la guanine 5' pour se reformer avec la troisième guanine : le déplacement chimique sur G_5 : H_8 varie peu au cours du temps. Cette étude a donc permis de préciser le comportement d'un adduit platiné sur la séquence G*G*G, d'explorer les possibilités de migration de l'adduit et d'améliorer la description du cisplatine dans le champ de force parm 98.

L'action anti-tumorale du cisplatine réside dans la capacité du dommage à recruter les protéines cellulaires de réparation et les facteurs de transcription. Pour décrire le comportement dynamique de la reconnaissance ADN-platiné/protéine, nous avons tout d'abord étudié les mécanismes de reconnaissance spécifique entre le facteur de transcription LEF-I (Lymphoid Enhanced-binding Factor - I) appartenant à la famille des HMG protéines (High Mobility Group). La simulation a révélé le rôle d'interactions électrostatiques entre le squelette de l'ADN et certains résidus de la protéine. Des contacts spécifiques de séquence ont permis de valider la reconnaissance séquence-spécifique de la LEF-I pour sa séquence cible. Durant la simulation de dynamique moléculaire, la flexibilité de la méthionine qui s'intercale entre les adénines désempilées dans le complexe ADN/protéine a créé une poche hydrophile qui est remplie par une, puis deux, puis une molécule d'eau. La stabilisation de l'intercalation par cette molécule d'eau semble être une caractéristique de l'affinité de reconnaissance.

Cette première simulation aurait dû être complétée par une simulation de la protéine avec la séquence platinée mais cette simulation n'a pu être réalisée. Cependant, les études d'affinité sur gel effectuées par le Dr. Drumm ont permis de valider l'hypothèse que le remplacement de la séquence AA où vient s'intercaler la méthionine par G*G* conservait l'affinité de la protéine pour la séquence platinée. Ces résultats expérimentaux suggèrent que le mode de reconnaissance n'est pas altéré et pourrait correspondre au comportement observé en simulation de dynamique moléculaire pour la séquence native.

La structure d'un ADN courbé par le cisplatine et la reconnaissance plausible par une protéine de type HMG ont permis de donner les premières pistes pour la description moléculaire des mécanismes initiant la cascade de réactions cellulaires, responsable de l'effet anti tumoral.

Le spectre d'activité forte du cisplatine étant restreint, le pyrazolate bisplatine a été conçu *de novo* et synthétisé par l'équipe du professeur Reedijk pour imposer comme le cisplatine une géométrie déterminée. La déformation présumée de ce complexe était *a priori* moins importante que celle du cisplatine. La combinaison de l'analyse RMN et de la simulation de dynamique moléculaire a validé l'hypothèse d'une déformation différente : le coude prononcé induit par le cisplatine sur son ADN cible est éliminé, mais le déroulement local d'environ 10° est conservé. L'activité cytotoxique du pyrazolato-bisplatine et l'analyse expérimentale indiquant une déformation différente de celle du cisplatine suggèrent que la voie métabolique utilisée par le pyrazolate bisplatine est différente de celle du cisplatine, ce qui permet d'envisager un spectre d'activité thérapeutique différent.

Les résultats expérimentaux ont permis de valider et d'améliorer la représentation des composés platinés dans le champ de force parm 98. Ce travail est poursuivi par le Dr. D. Djuranovic.

L'ensemble de ces travaux a permis une meilleure compréhension de la dynamique des adduits platinés et des modes de reconnaissance plausibles entre la déformation engendrée et la protéine qui vient s'y fixer.

Une meilleure description des composés platinés dans le champ de force de la suite logicielle AMBER est désormais disponible (et présentée en détail dans ce manuscrit).

Ce travail de thèse n'est cependant pas suffisant pour expliquer le tropisme du cisplatine et son efficacité si importante sur certains types de cancers. Nombre des cancers ciblés par le cisplatine sont situés dans les voies exposées du corps humain : système respiratoire et uro-génital. Pour nombre de ces cancers, l'origine est encore inconnue.

Selon Rosenberg, il est possible que l'efficacité du cisplatine vienne de la dérégulation de protéines régulatrices, ce qui entraînerait une réaction immunitaire. L'origine (et la validité) de cette réaction immunitaire n'est pas non plus élucidée mais pourrait provenir de l'activation de virus endogènes (présents sous forme de proto-oncogènes dans le génome humain), en particuliers d'adénovirus. Au moment de la rédaction de ce manuscrit, la structure et la séquence de nombreuses protéines virales est encore manquante, il n'est donc pas encore possible de valider/infirmier cette hypothèse.

Bibliographie

Bibliographie

- [1] Ahn H., Lee E., Kim K., and Lee C. Effect of glutathione and its related enzymes on chemosensitivity of renal cell carcinoma and bladder carcinoma cell lines. *J Urol*, 151(1), p. 263–7, 1994.
- [2] Alberts D.S. *Platinum and other metal coordination compounds in cancer chemotherapy 2*, pages 303–309. Plenum Press, New York, London, 1996.
- [3] Altona C. and Sundaralingam M. Conformational analysis of the sugar ring in nucleosides and nucleotides. a new description using the concept of pseudorotation. *J. Am. Chem. Soc.*, 94, p. 8205–8212, 1972.
- [4] Ano S. O., Kuklenyik Z., and Marzilli L.G. Structure and dynamics of pt anticancer drug adducts from nucleotides to oligonucleotides as revealed by NMR methods. In Lippert B., editor, *Cisplatin : Chemistry and Biochemistry of a Leading Anticancer Drug*, pages 247–291. Verlag Helvetica Chimica Acta, Zürich, 1999.
- [5] Arnott S., Chandrasekaran R., Hall I., and Puigjaner L. Heteronomous DNA. *Nucleic Acids Res.*, 11, p. 4141–4155, 1983.
- [6] Arora N. and Jayaram B. Strength of hydrogen bonds in *alpha*-helices. *J. Comp. Chem.*, 18, p. 1245–1252, 1997.
- [7] Auffinger P. and Westhof E. Simulations of the molecular dynamics of nucleic acids. *Curr. Opin. Struct. Biol.*, 8(2), p. 227–236, 1998.
- [8] Ban N., Nissen P., Hansen J., Moore P., and Steitz T. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science*, 289, p. 905–920, 2000.
- [9] Bancroft M., Williams L., Rich A., and Egli M. The low temperature crystal structure of the pure spermine form of Z-DNA reveals binding of a spermine molecule in the minor groove. *Biochemistry*, 33, p. 1073–1078, 1994.
- [10] Barnard C.F.J. *Plat. Med. Rev.*, 33, p. 162, 1989.

- [11] Basolo F., Gray H.B., and Pearson R.G. Mechanism of substitution reactions of complexes ions. XVII. Rates of reaction of some Platinum(II) and Palladium(II) complexes with pyridine. *J. Am. Chem. Soc.*, 82, p. 4200–4203, 1960.
- [12] Behe M. and Felsenfeld G. Effects of methylation on a synthetic polynucleotide : the B-Z transition in poly(dG-m⁵dC). *Proc. Natl. Acad. Sci. USA*, 78, p. 1619–1623, 1981.
- [13] Behrens B.C., Hamilton T.C., Masuda H., Grotzinger K.R., Whang-Peng J., Louie K.G., Knutsen T., McKoy W.M., Young R.C., and Ozols R.F. Characterization of a cis-diamminedichloroplatinum(II)-resistant human ovarian cancer cell line and its use in evaluation of platinum analogues. *Cancer Res.*, 47(2), p. 414–8, 1987.
- [14] Bellon S., Coleman J., and Lippard S. DNA unwinding produced by site-specific intrastrand cross-links of the antitumor drug cis-diamminedichloroplatinum(II). *Biochemistry*, 30(32), p. 8026–8035, 1991.
- [15] Bellon S. and Lippard S. Bending studies of DNA site-specifically modified by cis-platin, trans-diamminedichloroplatinum(II) and cis-[Pt(NH₃)₂(N3-cytosine)Cl]⁺. *Biophys. Chem.*, 35(2-3), p. 179–188, 1990.
- [16] Berendsen H. J. C., Postma J. P. M., Gunsteren W. F., Nola A. Di, and Haak J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81, p. 3684–3690, 1984.
- [17] Berman H., Bhat T., and Bourne P. The protein data bank and the challenge of structural genomics. *Nat. Struct. Biol.*, 7 (11), p. 957–959, 2000.
- [18] Berman H., Olson W., Beveridge D., Westbrook J., Gelbin A., Demeny T., Hsieh S., Srinivasan A., and Schneider B. The nucleic acid database : a comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.*, 63, p. 751–759, 1992.
- [19] Bernstein F., Koetzle T., Williams G., Meyer E. J., Brice M., Rodgers J., Kennard O., Shimanouchi T., and Tasumi M. The protein data bank : a computer-based archival file for macromolecular structure. *J. Mol. Biol.*, 112(3), p. 535–542, 1977.
- [20] Bohr V.A. Gene specific DNA repair. *Carcinogenesis*, 12(11), p. 1983–92, 1991.
- [21] Bradley L., Yarema K., Lippard S., and Essigmann J. Mutagenicity and genotoxicity of the major DNA adduct of the antitumor drug cis-diamminedichloroplatinum(II). *Biochemistry*, 32(3), p. 982–988, 1993.

- [22] Bruhn S L., Toney J. H., and Lippard S. J. Biological processing of DNA modified by platinum compounds. *Prog. Inorg. Chem.*, 38, p. 477–516, 1990. SL Bruhn, Mit, Dept Chem, Cambridge, MA 02139 edited by Stephen J. Lippard; John Wiley & Sons, Inc.
- [23] Bubley G.J., Ashburner B.P., and Teicher B.A. Spectrum of cis-diamminedichloroplatinum(II) induced mutations in a shuttle vector propagated in human cells. *Molec. Carcin.*, 4(5), p. 397–406, 1991.
- [24] Burnouf D., Gauthier C., Chottard J., and Fuchs R. Single d(ApG)/cis-diamminedichloroplatinum(II) adduct-induced mutagenesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, 87(16), p. 6087–6091, 1990.
- [25] Cantor C., Warshaw M., and Shapiro H. Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleolides. *Biopolymers*, 9(9), p. 1059 – 1077, 1970.
- [26] Case D. A., Pearlman D. A., Caldwell J. W., Cheatham III T. E., Ross W. S., Simmerling C. L., Darden T. A., K. M. Merz Jr., Stanton R. V., Cheng A. L., Vincent J. J., Crowley M., Tsui V., Radmer R. J., Duan Y., Pitera J., Massova I., Seibel G. L., Singh U. C., Weiner P. K., and Kollman P. A. AMBER 6, 1999.
- [27] Cech T., Zaug A., and Grabowski P. In vitro splicing of the ribosomal RNA precursor of tetrahymena : involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell*, 27, p. 487–496, 1981.
- [28] Celda B., Widmer H., Leupin W., Chaqin W., Denny W., and Wüthrich K. Conformational studies of d-(AAAAATTTTT)₂ using constraints from nuclear overhauser effects and from quantitative analysis of the cross-peak fine structures in two-dimensional 1h nuclear magnetic resonance spectra. *Biochemistry*, 28, p. 1462–1471, 1989.
- [29] Cheatham III T. E., Cieplak P., and Kollman P. A. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.*, 16(4), p. 845–862, 1999.
- [30] Cheatham III T. E. and Kollman P. A. Observation of the A-DNA to B-DNA transition during unrestrained molecular dynamics in aqueous solution. *J. Mol. Biol.*, 259(3), p. 434–444, 1996.
- [31] Cheatham III T. E. and Kollman P. A. Molecular dynamics simulation of nucleic acids. *Ann. Rev. Phy. Chem.*, 51, p. 435–471, 2000.

- [32] Cheatham III T. E. and Young M. Molecular dynamics simulation of nucleic acids : Successes, limitations, and promise. *Biopolymers*, 56(4), p. 232–256, 2000.
- [33] Cheatham III T.E. and Brooks B.R. Recent advances in molecular dynamics simulation towards realistic representation of biomolecules in solution. *Theor. Chem. Acc.*, 99, p. 279–288, 1998.
- [34] Cheatham III T.E. and Kollman P. A. Molecular dynamics simulations highlight the structural differences among DNA :DNA, RNA :RNA and DNA :RNA hybrid duplexes. *J. Am. Chem. Soc.*, 119, p. 4805–4825, 1997.
- [35] Cheng J., Chou S., Salazar M., and Reid B. Solution structure of [d(GCGTATACGC)]₂. *J. Mol. Biol.*, 228, p. 118–137, 1992.
- [36] Chevrier B., Dock A., Hartmann B., Leng M., Moras D., Thuong M., and Westhof E. Solvation of the left-handed hexamer d(⁵Br C-G-⁵Br C-G-⁵Br C-G) in crystals grown at two temperatures. *J. Mol. Biol.*, 188, p. 709–719, 1986.
- [37] Chow C., Whitehead J., and Lippard S. HMG domain proteins induce sharp bends in cisplatin-modified DNA. *Biochemistry*, 33(50), p. 15124–15130, 1994.
- [38] Chval Z. and Sip M. Force field for platinum binding to adenine and guanine taking into account flexibility of nucleic acids bases. *J. Phys. Chem.*, B 102, p. 1659–1661, 1998.
- [39] Ciccarelli R., Solomon M., Varshavsky A., and Lippard S. In vivo effects of cis- and trans-diamminedichloroplatinum(II) on SV40 chromosomes : differential repair, DNA-protein cross-linking, and inhibition of replication. *Biochemistry*, 24(26), p. 7533–7540, 1985.
- [40] Cizeau J., Decoville M., Leng M., and Locker D. Deletions induced in the white and vermilion genes of drosophila melanogaster by the antitumor drug cis-dichlorodiammineplatinum(II). *Mutat. Res.*, 311(1), p. 31–38, 1994.
- [41] Cleare M.J. and Hoeschele J.D. *Plt. Med. Rev.*, 17, p. 3, 1973.
- [42] Cleare M.J. and Hoeschele J.D. Studies on the antitumour activity of group VIII transition metal complexes. Part. 1. Platinum II complexes. *Bioinorg. Chem.*, 2, p. 187, 1973.
- [43] Comess K., Burstyn J., Essigmann J., and Lippard S. Replication inhibition and translesion synthesis on templates containing site-specifically placed cis-diamminedichloroplatinum(II) DNA adducts. *Biochemistry*, 31(16), p. 3975–3990, 1992.

- [44] Corda Y., Anin M., Leng M., and Job D. RNA polymerases react differently at d(ApG) and d(GpG) adducts in DNA modified by cis-diamminedichloroplatinum(II). *Biochemistry*, 31(7), p. 1904–1908, 1992.
- [45] Corda Y., Job C., Anin M., Leng M., and Job D. Transcription by eucaryotic and procaryotic RNA polymerases of DNA modified at a d(GG) or a d(AG) site by the antitumor drug cis-diamminedichloroplatinum(II). *Biochemistry*, 30(1), p. 222–230, 1991.
- [46] Corda Y., Job C., Anin M., Leng M., and Job D. Spectrum of DNA-platinum adduct recognition by prokaryotic and eukaryotic DNA-dependent RNA polymerases. *Biochemistry*, 32(33), p. 8582–8588, 1993.
- [47] Cornell W. D., Cieplak P., Bayly C. I., Gould I. R., K. M. Merz Jr., Ferguson D. M., Spellmeyer D. C., Fox T., Caldwell J. W., and Kollman P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, 117, p. 5179–5197, 1995.
- [48] Darden T. A., York D., and Pedersen L. G. Particle mesh ewald : an $n\log(n)$ method for ewald sums in large systems. *J. Chem. Phys.*, 98(12), p. 10089–10092, 1993.
- [49] Hartog J. H. J., Altona C., Boom J. H., Marel G. A., Haasnoot C. A. G., and Reedijk J. Cis-diamminedichloroplatinum(II) induced distortion of a single and double strand deoxydecanucleosidenonaphosphate studied by NMR. *J. Biomol. Struct. Dyn.*, 2(6), p. 1137–1155, 1985.
- [50] Dickerson R. DNA structure from A to Z. *Methods Enzymol.*, 211, p. 67–111, 1992.
- [51] Dickerson R.E., Bansal M., Calladine C.R., Diekman S., Hunter W.N., and Kennard O. Definition and nomenclature of nucleic acid structure parameters. *J. Mol. Biol.*, 205, p. 787–791, 1989.
- [52] Dijt F.J., Fichtinger-Schepman A.M.J., Berends F., and Reedijk J. Formation and repair of cisplatin-induced adducts to DNA in cultured normal and repair deficient human fibroblasts. *Cancer Res.*, 48, p. 6058–6062, 1988.
- [53] Doktycz M., Benight A., and Sheardy R. Energetics of B-Z junction formation in a sixteen base-pair duplex DNA. *J. Mol. Biol.*, 212, p. 3–6, 1990.
- [54] Drew H. and Dickerson R. Structure of a B-DNA dodecamer. III. Geometry of hydration. *J. Mol. Biol.*, 151(3), p. 535–556, 1981.

- [55] Dunham S. and Lippard S. DNA sequence context and protein composition modulate HMG-domain protein recognition of cisplatin-modified DNA. *Biochemistry*, 36(38), p. 11428–11436, 1997.
- [56] Dunham S. U., Dunham S. U., Turner C. J., and Lippard S. J. Solution Structure of a DNA Duplex Containing a Nitroxide Spin-Labeled Platinum d(GpG) Intrastrand Cross-Link Refined with NMR-Derived Long-Range Electron-Proton Distance Restraints. *J. Am. Chem. Soc.*, 120(22), p. 5395–5406, 1998.
- [57] Eastman A. The formation, isolation, and characterisation of DNA adducts produced by anticancer platinum complexes. *Pharmacol. Ther.*, 34, p. 155–166, 1987.
- [58] Elizondo-Riojas M. and Kozelka J. Asymmetrical complexes [pt(n,n-dmen)(dguo)l]((2)+) (n,n-dmen = n,n-dimethylethylenediamine; l = h₂o, cl-, dguo) : H8 nuclei experience an orientation-dependent deshielding effect attributable to the paramagnetic anisotropy of the platinum atom. *Inorganica Chimica Acta*, 297(1-2), p. 417–420, 2000.
- [59] Elizondo-Riojas M. and Kozelka J. Unrestrained 5 ns molecular dynamics simulation of a cisplatin- DNA 1,2-GG adduct provides a rationale for the NMR features and reveals increased conformational flexibility at the platinum binding site. *J. Mol. Biol.*, 314(5), p. 1227–1243, 2001.
- [60] Elizondo-Riojas M.-A., Gonnet F., Augé-Barrere-Mazouat P., Allain F., Bergès J., Attias R., Chottard J.-C., and Kozelka J. Molecular modeling of platinum complexes with oligonucleotides : Methodological lessons and structural insights. In Banci L. and Comba P., editors, *Molecular Modeling and Dynamics of Bioinorganic Systems*, pages 131–160. Kluwer academic publishers, Dordrecht, 1997.
- [61] Ewald P. Due berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der Physik*, 64, p. 253–287, 1921.
- [62] Farrell N., Qu Y., Feng L., and Houten B. Van. Comparison of chemical reactivity, cytotoxicity, interstrand cross-linking and DNA sequence specificity of bis(platinum) complexes containing monodentate or bidentate coordination spheres with their monomeric analogues. *Biochemistry*, 29(41), p. 9522–9531, 1990.
- [63] Farrell N., Qu Y., and Hacker M. Cytotoxicity and antitumor activity of bis(platinum) complexes. a novel class of platinum complexes active in cell

- lines resistant to both cisplatin and 1,2-diaminocyclohexane complexes. *J. Med. Chem.*, 33(8), p. 2179–2184, 1990.
- [64] Feig M. and Pettitt M. Sodium and chlorine ions as part of the DNA solvation shell. *Biophys. J.*, 77, p. 1769–1781, 1999.
- [65] Fouchet M.-H., Guittet E., Cognet J. A. H., Kozelka J., Gauthier C., Bret M. Le, Zimmermann K., and Chottard J.-C. Structure of a nonanucleotide duplex cross-linked by cisplatin at an ApG sequence. *J. Biol. Inorg. Chem.*, 2, p. 83–92, 1997.
- [66] Franklin R. and Gosling R. Molecular structure of nucleic acids : molecular configuration in sodium thymonucleate. *Nature*, 171, p. 737–741, 1953.
- [67] Fratini A.V., Kopka M.L., Drew H.R., and Dickerson R.E. Reversible bending and helix geometry in a B-DNA dodecamer : CGCGAATTBrCGCG. *J. Biol. Chem.*, 257(24), p. 14686–14707, 1982.
- [68] Fuertes M.A., Alonso C., and Perez J.M. Biochemical modulation of Cisplatin mechanisms of action : enhancement of antitumor activity and circumvention of drug resistance. *Chem. Rev.*, 103(3), p. 645–62, 2003.
- [69] Fuertesa M.A., Castillab J., Alonsoa C., and Perez J.M. Cisplatin biochemical mechanism of action : from cytotoxicity to induction of cell death through interconnections between apoptotic and necrotic pathways. *Curr. Med. Chem.*, 10(3), p. 257–66, 2003.
- [70] Fujii S., Wang A., Marel G., Boom J., and Rich A. Molecular structure of (m⁵dCdG)₃ : the role of the methyl group on 5-methyl cytosine in stabilizing Z-DNA. *Nucleic Acids Res.*, 23, p. 7879–7891, 1982.
- [71] Gelasco A. and Lippard S. NMR solution structure of a DNA dodecamer duplex containing a cis-diammineplatinum(II) d(GpG) intrastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochemistry*, 37(26), p. 9230–9239, 1998.
- [72] Gessner R., Quigley G., and Egli M. Comparative studies of high resolution Z-DNA crystal structures : common hydration patterns of alternating dC-dG. *J. Mol. Biol.*, 236, p. 1154–1168, 1994.
- [73] Giessner-Prettre C. and Pullman B. Ring-current effects in the NMR of nucleic acids : A graphical approach. *Biopolymers*, 15, p. 2277–2286, 1976.
- [74] Godwin A.K., Meister A., O'Dwyer P.J., Huang C.S., Hamilton T.C., and Anderson M.E. High resistance to cisplatin in human ovarian cancer cell lines is

- associated with marked increase of glutathione synthesis. *Proc Natl Acad Sci U S A*, 89(7), p. 3070–4, 1992.
- [75] Gorenstein D. ^{31}P NMR of DNA. *Methods Enzymol.*, 211, p. 254–286, 1992.
- [76] Gupta G., Bansal M., and Sasisekharan V. Conformational flexibility of DNA : polymorphism and handedness. *Proc. Natl. Acad. Sci. USA*, 77, p. 6486–6490, 1980.
- [77] Guzikevich-Guerstein G. and Shakked Z. A novel form of the DNA double helix imposed on the TATA-box by the TATA-binding protein. *Nat. Struct. Biol.*, 3, p. 32–37, 1996.
- [78] Guéron M. and Desmaret J. A simple explanation of the electrostatics of the B-to-Z transition of DNA. *Proc. Natl. Acad. Sci. USA*, 89, p. 5740–5743, 1992.
- [79] Hanawalt P.C. Transcription-coupled repair and human disease. *Science*, 266(5193), p. 1957–1958, 1994.
- [80] Haran T., Shakked Z., Wang A., and Rich A. The crystal structure of d(CCCCGGGG) : a new A-form variant with an extended backbone conformation. *J. Biomol. Struct. Dyn.*, 5, p. 199–217, 1987.
- [81] Hartmann B. and Lavery R. DNA structural forms. *Q. Rev. Biophys.*, 29(4), p. 309–368, 1996.
- [82] Harvey S. DNA structural dynamics : longitudinal breathing as a possible mechanism for the B-Z transition. *Nucleic Acids Res.*, 11, p. 4867–4873, 1983.
- [83] Harvey S. C., Tan R. K.-Z., and Cheatham III T. E. The flying ice cube : Velocity rescaling in molecular dynamics leads to violation of energy equipartition. *J. Comput. Chem.*, 19(7), p. 726–740, 1998.
- [84] Haworth I., Rodger A., and Richards W. A possible mechanism for the B to Z transition. *J. Biomol. Struct. Dyn.*, 10, p. 195–211, 1992.
- [85] Heiger-bernays W.J., Essigman J.M., and Lippard S.J. Effect of the antitumor drug cis-diamminedichloroplatinum(II) and related platinum complexes on eukaryotic DNA replication. *Biochemistry*, 29(36), p. 8461–8466, 1990.
- [86] Heinemann U., Alings C., and Hahn M. Crystallographic studies of DNA helix structure. *Biophys. Chem.*, 50, p. 157–167, 1994.

- [87] Herman F., Kozelka J., Stoven V., Guittet E., Girault J., HuyNH-DiNH T., Igolen J., Lallemand J., and Chottard J. A d(GpG)-platinated decanucleotide duplex is kinked. an extended NMR and molecular mechanics study. *Eur. J. Biochem.*, 194(1), p. 119–133, 1990.
- [88] Holler E., Bauer R., and Bernges F. Monofunctional DNA-platinum(II) adducts block frequently DNA polymerases. *Nucleic Acids Res.*, 20(9), p. 2307–2312, 1992.
- [89] Hollis L., Sundquist W., Burstyn J., Heiger-Bernays W., Bellon S., Ahmed K., Amundsen A., Stern E., and Lippard S. Mechanistic studies of a novel class of trisubstituted platinum(II) antitumor agents. *Cancer Res.*, 51(7), p. 1866–1875, 1991.
- [90] Hollis S.T., Amundsen A.R., and Stern E.W. Chemical and biological properties of a new series of *cis*-diammineplatinum(II) antitumor agents containing three nitrogen donors : *cis*-[Pt(NH₃)₃(N-donor)Cl]⁺. *J. Med. Chem.*, 32, p. 128–136, 1989.
- [91] Hoogsteen K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Cryst.*, 16, p. 907–916, 1963.
- [92] <http://www.ccdc.cam.ac.uk> . Cambridge crystallographic data centre. 2004.
- [93] Hua N., Marel G., Boom J., and Feignon J. Non contiguous regions of Z-DNA in a DNA dodecamer. *Nucleic Acids Res.*, 17, p. 7923–8044, 1989.
- [94] Huang H., Zhu L., Reid B. R., Drobny G. P., and Hopkins P. B. Solution structure of a cisplatin-induced DNA interstrand cross-link. *Science*, 270, p. 1842–1845, 1995.
- [95] Humphrey W., Dalke A., and Schulten K. VMD - Visual Molecular Dynamics. *J. Molec. Graphics*, 14(1), p. 33–38, 1996.
- [96] James T. Assessment of quality of derived macromolecular structures. *Methods Enzymol.*, 239, p. 416–439, 1994.
- [97] Johnson S.W., Perez R.P., Godwin A.K., Yeung A.T., Handel L.M., Ozols R.F., and Hamilton T.C. Role of platinum-DNA adduct formation and removal in cisplatin resistance in human ovarian cancer cell lines. *Biochem. Pharmacol.*, 47(4), p. 689–97, 1994.

- [98] Johnson S.W., Swiggard P.A., Handel L.M., Brennan J.M., Godwin A.K., Ozols R.F., and Hamilton T.C. Relationship between platinum-DNA adduct formation and removal and cisplatin cytotoxicity in cisplatin-sensitive and -resistant human ovarian cancer cells. *Cancer Res.*, 54(22), p. 5911–6, 1994.
- [99] Johnston B. Generation and detection of Z-DNA. *Methods Enzymol.*, 211, p. 127–159, 1992.
- [100] Jorgensen W. L., Chandrasekhar J., Madura J. D., Impey R. W., and Klein M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79(2), p. 926–935, 1983.
- [101] Kasparkova J. and Brabec V. Recognition of DNA interstrand cross-links of cis-diamminedichloroplatinum(II) and its trans isomer by DNA-binding proteins. *Biochemistry*, 34(38), p. 12379–12387, 1995.
- [102] Kasparkova J., Zehnulova J., Farrell N., and Brabec V. DNA interstrand cross-links of the novel antitumor trinuclear platinum complex BBR3464 - conformation, recognition by high mobility group domain proteins, and nucleotide excision. *J. Biol. Chem.*, 277(50), p. 48076–48086, 2002.
- [103] Kelland L.R. Preclinical perspectives on platinum resistance. *Drugs*, 59 Suppl 4, p. 1–8 ; discussion 37–8, 2000.
- [104] Kennard O., Cruse W., and Nachman J. Ordered-water structure in an A-DNA octamer at 1.7 Å resolution. *J. Biomol. Struct. Dyn.*, 3, p. 623–647, 1986.
- [105] Klump H., Schmid E., and Wosgien M. Energetics of Z-DNA formation in poly d(A-T), poly d(G-C) and poly d(A-C).poly d(G-T). *Nucleic Acids Res.*, 21, p. 2343–2348, 1993.
- [106] Kozelka J. Computational studies on platinum antitumor complexes and their adducts with nucleic acids constituents. In Lippert B., editor, *Cisplatin : Chemistry and Biochemistry of a Leading Anticancer Drug*, pages 537–556. Verlag Helvetica Chimica Acta, Zürich, 1999.
- [107] Kozelka J. and Chottard J.-C. How does cisplatin alter DNA structure ? a molecular mechanics study on double-stranded oligonucleotides. *Biophys. Chem.*, 35(2-3), p. 165–178, 1990.
- [108] Kozelka J., Savinelli R., Berthier G., Flament J. P., and Lavery R. Force field for platinum binding to adenine. *J. Comput. Chem.*, 14(1), p. 45–53, 1993.

- [109] Kyoguku Y., Lord R., and Rich A. An infrared study of hydrogen bonding between adenine and uracil derivatives in chloroform solution. *J. Am. Chem. Soc.*, 89, p. 496–504, 1967.
- [110] Lai G.M., Ozols R.F., Smyth J.F., Young R.C., and Hamilton T.C. Enhanced DNA repair and resistance to cisplatin in human ovarian cancer. *Biochem Pharmacol*, 37(24), p. 4597–600, 1988.
- [111] Lavery R. *Advances in Computational Biology*. Jai Press Inc., 1994.
- [112] Lavery R. and Sklenar H. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.*, 6(1), p. 63–91, 1988.
- [113] Lavery R. and Sklenar H. Defining the structure of irregular nucleic acids : conventions and principles. *J. Biomol. Struct. Dyn.*, 6(4), p. 655–67, 1989.
- [114] Lavery R. and Zakrzewska K. *Oxford handbook of nucleic acid structure*. Oxford University Press, 1997.
- [115] Lavery R., Zakrzewska K., and Sklenar H. Jumna (junction minimisation of nucleic acids). *Comp. Phys. Comm.*, 91, p. 135–158, 1995.
- [116] Lebrun A. and Lavery R. Unusual DNA conformations. *Curr. Opin. Struct. Biol.*, 7(3), p. 348–354, 1997.
- [117] Leng M. Left-handed Z-DNA. *Biochim. Biophys. Acta*, 825(4), p. 339–344, 1985.
- [118] Lippert B, Raudaschl G, Lock CJL, and Pilon P. 'Real' model compounds for intrastrand cross-linking of two guanine bases by cisplatin : Crystal structures of cis-diamminebis(9-ethylguanine-N₇)platinum(II) dichloride trihydrate, [Pt(NH₃)₂(C₇H₉N₅O)₂]Cl₂ · 3H₂O, and cis-diamminebis(9-ethylguanine-N₇) platinum(II) sesquichloride hemibicarbonate sesquihydrate, [Pt(NH₃)₂(C₇H₉N₅O)₂]Cl_{1.5}(HCO₃)_{0.5} · 1.5H₂O. *Inorg. Chim. Acta*, 93(1), p. 43–50, 1984.
- [119] Liquiers J., Taillandier E., Peticolas W., and Thomas G. The infrared and raman spectra of the duplex of d(GGTATACC) in the crystal show bands due to both the A-form and the B-form of DNA. *J. Biomol. Struct. Dyn.*, 8, p. 295–302, 1990.
- [120] Loehrer P. J. and Einhorn L. H. Cisplatin. *Ann. Intern. Med.*, 100, p. 704–713, 1984.

- [121] Loprete D. and Hartman K. Conditions for the stability of the B, C, and Z structural forms of poly(dG-dC) in the presence of lithium, potassium, magnesium, calcium and zinc cations. *Biochemistry*, 32, p. 4077–4082, 1993.
- [122] Love J., Li X., Case D., Giese K., Grosschedl R., and Wright P. Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature*, 376, p. 791–795, 31 August 1995.
- [123] Malfoy B., Rousseau N., and Leng M. Interaction between antibodies to Z-form deoxyribonucleic acid and double-stranded polynucleotides. *Biochemistry*, 21(22), p. 5463–5467, 1982.
- [124] Malinge J., Perez C., and Leng M. Base sequence-independent distortions induced by interstrand cross-links in cis-diamminedichloroplatinum (II)-modified DNA. *Nucleic Acids Res.*, 22(19), p. 3834–3839, 1994.
- [125] Marcelis A.T.M., Hartog J.H.J., Marel G.A. Van, Wille G., and Reedijk J. Interaction of platinum compounds with short oligodeoxynucleotides containing guanine and cytosine. *Eur. J. Biochem.*, 135(2), p. 343–349, 15 September 1983.
- [126] Marky L. and Breslauer K. Calorimetric determination of base-stacking enthalpies in double helical DNA molecules. *Biopolymers*, 21, p. 2185–2194, 1982.
- [127] Marzec C. and Day L. An exact description of five-membered ring configurations. I. parameterization via an amplitude S, an angle gamma, the pseudorotation amplitude q and phase angle P, and the bond lengths. *J. Biomol. Struct. Dyn.*, 10(6), p. 1091–1123, May 1993.
- [128] Marzec C. and Day L. An exact description of five-membered ring configurations. II. applications to furanose rings in DNA and RNA, analysis of errors, and bond angle bending energy. *J. Biomol. Struct. Dyn.*, 10(6), p. 1125–1154, May 1993.
- [129] Marzilli L., Saad J., Kuklenyik Z., Keating K., and Xu Y. Relationship of solution and protein-bound structures of DNA duplexes with the major intrastrand cross-link lesions formed on cisplatin binding to DNA. *J. Am. Chem. Soc.*, 123(12), p. 2764–2770, 2001.
- [130] Masuda H., Ozols R.F., Lai G.M., Fojo A., Rothenberg M., and Hamilton T.C. Increased DNA repair as a mechanism of acquired resistance to cis-diamminedichloroplatinum (II) in human ovarian cancer cell lines. *Cancer Res*, 48(20), p. 5713–6, 1988.

- [131] Masuda H., Tanaka T., Matsuda H., and Kusaba I. Increased removal of DNA-bound platinum in a human ovarian cancer cell line resistant to cis-diamminedichloroplatinum(II). *Cancer Res.*, 50(6), p. 1863–6, 1990.
- [132] Mauffret O., Hartmann B., Convert O., Lavery R., and Fermandjian S. The fine structure of two DNA dodecamers containing the cAMP responsive element sequence and its inverse. nuclear magnetic resonance and molecular simulation studies. *J. Mol. Biol.*, 227(3), p. 852–875, 5 October 1992.
- [133] McA’Nulty M., Whitehead J., and Lippard S. Binding of Ixrl, a yeast HMG-domain protein, to cisplatin-DNA adducts in vitro and in vivo. *Biochemistry*, 35(19), p. 6089–6099, 14 May 1996.
- [134] Metzkurshel U., Kurscehl E., Niederle N., and Aulbert E. Investigations on the acute and chronic nephrotoxicity of the new platinum analogue carboplatin. *J. Cancer Res. Clin. Oncol.*, 116, p. 203–206, 1990.
- [135] Minchekova L., Schyolkina A., Chernov B., and Ivanov V. CC/GG contacts facilitate the B to A transition of DNA in solution. *J. Biomol. Struct. Dyn.*, 4, p. 463–476, 1986.
- [136] Mistry P., Kelland L.R., Abel G., Sidhar S., and Harrap K.R. The relationships between glutathione, glutathione-S-transferase and cytotoxicity of platinum drugs and melphalan in eight human ovarian carcinoma cell lines. *Br. J. Cancer*, 64(2), p. 215–20, 1991.
- [137] Monjardet-Bas V., Chottard J., and Kozelka J. Fast interstrand cross-linking of cisplatin-DNA monoadducts compared with intrastrand chelation : A kinetic study using hairpin-stabilized duplex oligonucleotides. *Chemistry-a European Journal*, 8(5), p. 1144–1150, 2002.
- [138] Newmark R. and Cantor C. Nuclear magnetic resonance study of the interaction of guanosine and cytidine in dimethyl sulfoxide. *J. Am. Chem. Soc.*, 90, p. 5010–5017, 1968.
- [139] Ohndorf U. M., Rould M. A., He Q., Pabo C. O., and Lippard S. J. Basis for recognition of cisplatin-modified DNA by high-mobility-group proteins. *Nature*, 399, p. 708–712, 1999.
- [140] Paquet F., Perez C., Leng M., Lancelot G., and Malinge J. NMR solution structure of a DNA decamer containing an interstrand cross-link of the antitumor drug cis-diamminedichloroplatinum (II). *J. Biomol. Struct. Dyn.*, 14(1), p. 67–77, 1996.

- [141] Parker R.J., Eastman A., Bostick-Bruton F., and Reed E. Acquired cisplatin resistance in human ovarian cancer cells is associated with enhanced repair of cisplatin-DNA lesions and reduced drug accumulation. *J Clin Invest*, 87(3), p. 772–7, 1991.
- [142] Parkinson J. A., Chen Y., Guo Z., Berners-Price S. J., Brown T., and Sadler P. J. Sequence-dependent bending of DNA induced by cisplatin : NMR structures of an A.T rich 14-mer duplex. *Chem. Eur. J.*, 6(19), p. 3636–3644, 2000.
- [143] Patrick S.M., Henkels K.M., and Turchi J.J. High-Mobility Group 1 protein inhibits helicase catalyzed displacement of cisplatin-damaged DNA. *Biochim. Biophys. Acta*, 1354(3), p. 279–290, 1997.
- [144] Patrick S.M. and Turchi J.J. Human replication protein A preferentially binds cisplatin-damaged duplex DNA *in vitro*. *Biochemistry*, 37(24), p. 8808–8815, 1998.
- [145] Pavletich N. and Pabo C. Zinc finger-DNA recognition : crystal structure of a zif268-DNA complex at 2.1 Å. *Science*, 252, p. 809–817, 1991.
- [146] Perez R.P. Cellular and molecular determinants of cisplatin resistance. *Eur J Cancer*, 34(10), p. 1535–42, 1998.
- [147] Peyrone M. *Ann. Chem. Pharm.*, 51, p. 15, 1845.
- [148] Pil P. and Lippard S. Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. *Science*, 256(5054), p. 234–7, 1992.
- [149] Pinto A.L. and Lippard S.J. Sequence dependent termination in vitro DNA synthesis by cis- and trans- diamminedichloroplatinum(II). *Proc. Natl. Acad. Sci. USA*, 82, p. 4616–4619, 1985.
- [150] Pohl F. and Jovin T. Salt-induced cooperative conformational change of a synthetic DNA : equilibrium and kinetic studies with poly(dG-dC). *J. Mol. Biol.*, 67, p. 375–396, 1972.
- [151] Polak M., Plavec J., Trifonova A., Foldesi A., and Chattopadhyaya J. The change in the electronic character upon cisplatin binding to guanine nucleotide is transmitted to drive the conformation of the local sugar-phosphate backbone - a quantitative study. *J. Chem. Soc. - Perkin Transactions 1*, (19), p. 2835–2843, 1999.
- [152] Pratt W.B., Ruddon R.W., Ensminger W.D, and Maybaum J. *Section : 6. The anticancer drugs - Covalent DNA binding drugs - Platinum compounds*, pages 133–139. Oxford University Press, New york, Oxford, 1994.

- [153] Price J. and Cech T. Determinants of the 3' splice site for self-splicing of the tetrahymena pre-rRNA. *Genes Dev.*, 2, p. 1439–1447, 1998.
- [154] Privé G., Heinemann U., Chandrasegaran S., Kan L., Kopka M., and Dickerson R. Helix geometry, hydration, and G.A mismatch in a B-DNA decamer. *Science*, 238, p. 498–504, 23 October 1987.
- [155] Privé G., Yanagi K., and Dickerson R. Structure of the B-DNA decamer C-C-A-A-C-G-T-T-G-G and comparison with isomorphous decamers C-C-A-A-G-A-T-T-G-G and C-C-A-G-G-C-C-T-G-G. *J. Mol. Biol.*, 217(1), p. 177–199, 5 January 1991.
- [156] Pullman A. and Pullman B. Molecular electrostatic potential of the nucleic acids. *Quart. Rev. Biophys.*, 14(3), p. 289–380, 1981.
- [157] Rao S. and Kollman P. On the role of uniform and mixed sugar puckers in DNA double-helical structures. *J. Am. Chem. Soc.*, 107, p. 1611–1617, 1985.
- [158] Reedijk J., Fichtinger-Schepman A.M.J., Oosterom A.T., and Putte P. Platinum amine coordination compounds as anti-tumor drugs. Molecular aspects of the mechanism of action. *Struct. Bonding*, 67, p. 53–89, 1987.
- [159] Reich Z., Friedman P., Scolnik Y., Sussman J., and Minsky A. On the metastability of the left-handed DNA motifs. *Biochemistry*, 32, p. 2116–2119, 1993.
- [160] Rinkel L. J. and Altona C. Conformational analysis of the deoxyribofuranose ring in DNA by means of sums of proton-proton coupling constants : a graphical method. *J. Biomol. Struct. Dyn.*, 4(CHAP 7), p. 621–651, 1987.
- [161] Rosenberg B., Camp L. Van, Gumley E.B., and Thomson A.J. The inhibition of growth or cell division in *Escherichia Coli* by different ionic species of platinum(IV) complexes. *J. Biol. Chem.*, 242(6), p. 1347–1352, 1967.
- [162] Rosenberg B., Camp L. Van, and Krigas T. Inhibition of cell division in *Escherichia Coli* by electrolysis products from a platinum electrode. *Nature*, 205(4972), p. 698–699, 1965.
- [163] Rosenberg B., Camp L. Van, Trosko J.E., and Mansoui V.H. Platinum compounds : a new class of potent antitumor agents. *Nature*, 222, p. 385–386, 1969.
- [164] Ryckaert J.P., Ciccotti G., and Berendsen H.J.C. Numerical integraton of the cartesian equations of motion of a system with constraints : molecular dynamics of n-alkanes. *J. Comput. Phys.*, 23, p. 327–341, 1977.

- [165] Saenger W., Hunter W., and Kennard O. DNA conformation is determined by economics in the hydration of phosphate groups. *Nature*, 324, p. 385–388, 1986.
- [166] Sagui C. and Darden T.A. Molecular dynamics simulations of biomolecules : Long-range electrostatic effects. *Annu. Rev. Biophys. Biomol. Struct.*, 28, p. 155–179, 1999.
- [167] Salles B., Butour J.L., Lesca C., and Macquet J.P. cis-Pt(NH₃)₂Cl₂ and trans-Pt(NH₃)₂Cl₂ inhibit DNA synthesis in cultured L1210 leukemia cells. *Biochem. Biophys. Res. Comm.*, 112(2), p. 555–563, 1983.
- [168] Sanghani S. R. and Lavery R. Theoretical studies of DNA-rna hybrid conformations. *Nucleic Acids Res.*, 22(8), p. 1444–1449, 1994.
- [169] Scheek R. M., Boelens R., Russo N., Boom J. H., and Kaptein R. Sequential resonance assignments in 1H NMR spectra of oligonucleotides by two-dimensional NMR spectroscopy. *Biochemistry*, 23(23), p. 1371–1376, 1984.
- [170] Schöllhorn H., Raudaschl-Sieber G., Müller G., Thewalt U., and Lippert B. DNA-intrastrand guanine, guanine cross-linking by cisplatin : comparison of three model compounds with head-head orientation of the nucleobases. *J. Am. Chem. Soc.*, 107, p. 5932–5937, 1985.
- [171] Shakked Z. The influence of the environment on the DNA double helix. *Curr. Opin. Struct. Biol.*, 1, p. 446–451, 1991.
- [172] Sheardy R. and Winkle S. Temperature dependent CD and NMR studies on a synthetic oligonucleotide containing a B-Z junction at high salt. *Biochemistry*, 28, p. 720–725, 1989.
- [173] Shepard W., Cruse W., Fourme R., Fortelle E., and Prangé T. A zipper-like duplex in DNA : the crystal structure of d(GCGAAAGCT) at 2.1 Å resolution. *Structure*, 6(7), p. 849–861, 15 July 1998.
- [174] Shui X., McFail-Isom L., Hu G., and Williams L. The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry*, 37, p. 8341–8355, 1998.
- [175] Shui X., Sines C., McFail-Isom L., VanDerveer D., and Williams L. Structure of the potassium form of CGCGAATTCGCG : DNA deformation by electrostatic collapse around inorganic cations. *Biochemistry*, 37, p. 16877–16887, 1998.

- [176] Simmerling C., Elber R., and Zhang J. *MOIL-View - A Program for Visualization of Structure and Dynamics of Biomolecules and STO- A Program for Computing Stochastic Paths*, volume *Modelling of Biomolecular Structure and Mechanisms*, pages 241–265. Kluwer, Netherlands, 1995.
- [177] Sorenson C.M. and Eastman A. Influence of cis-diamminedichloroplatinum(II) on DNA synthesis and cell cycle : role of G2 arrest and DNA double-strand breaks. *Cancer Res.*, 48(23), p. 6703–6707, 1 December 1988.
- [178] Sorenson C.M. and Eastman A. Mechanism of cis-diamminedichloroplatinum(II) induced cytotoxicity : role of G2 arrest and DNA double-strand breaks. *Cancer Res.*, 48(16), p. 4484–4488, 15 August 1988.
- [179] Stofer E., Chipot C., and Lavery R. Free energy calculations of Watson-Crick base pairing in aqueous solution. *J. Am. Chem. Soc.*, 121(41), p. 9503–9508, 1999.
- [180] Suzuki M., Amano N., Kakinuma J., and Tateno M. Use of a 3D structure database for understanding sequence-dependent conformational aspects of DNA. *J. Mol. Biol.*, 274(3), p. 421–435, 5 December 1997.
- [181] Takahara P., Rosenzweig A., Frederick C., and Lippard S. Crystal structure of double-stranded DNA containing the major adduct of the anticancer drug cisplatin. *Nature*, 377(6550), p. 649–652, 1995.
- [182] Takahara P. M., Frederick C. A., and Lippard S. J. Crystal structure of the anticancer drug cisplatin bound to duplex DNA. *J. Am. Chem. Soc.*, 118(49), p. 12309–12321, 1996.
- [183] Takahara P. M., Frederick C. A., and Lippard S. J. Crystal structure of the anticancer drug cisplatin bound to duplex DNA. *J. Am. Chem. Soc.*, 119(20), p. 4795, 1997.
- [184] Takeuchi H., Hanamura N., and Harada I. Structural specificity of peptides in Z-DNA formation and energetics of the peptide induced B-Z transition of poly(dG-m⁵dC). *J. Mol. Biol.*, 236, p. 610–617, 1994.
- [185] Teletchea S., Hartmann B., and Kozelka J. Discrimination between BI and BII conformational substates of B-DNA based on sugar-base interproton distances. *J Biomol Struct Dyn*, 21(4), p. 489–94, 2004.
- [186] Timsit Y. and Moras D. Crystallization of DNA. *Methods Enzymol.*, 211, p. 409–429, 1992.

- [187] Tisne C., Hartmann B., and Delepierre M. NF-kappa B binding mechanism : a nuclear magnetic resonance and modeling study of a GGG \rightarrow CTC mutation. *Biochemistry*, 38(13), p. 3883–3894, 1999.
- [188] Treiber D.K., Zhai X., Jantzen H.M., and Essigmann J.M. Cisplatin-DNA adducts are molecular decoys for the ribosomal RNA transcription factor hUBF (human upstream binding factor). *Proc. Natl. Acad. Sci. USA*, 91(12), p. 5672–5676, 1994.
- [189] Trimmer E. E., Zamble D. B., Lippard S. J., and Essigman J. M. Human testis-determining factor SRY binds to the major DNA adduct of cisplatin and a putative target sequence with comparable affinities. *Biochemistry*, 37(1), p. 352–362, 6 January 1998.
- [190] Ven F. J. M. and Hilbers C. W. Resonance assignments of non-exchangeable protons in b type DNA oligomers, an overview. *Nucl. Acids Res.*, 16, p. 5713–5726, 1988.
- [191] Van Garderen C.J. and Van Houte L.P.A. The solution structure of a DNA duplex containing the cis-Pt(NH₃)₂[d(-GTG-)-N₇(G), N₇(G)] adduct, as determined with high-field NMR and molecular mechanics/dynamics. *Eur. J. Biochem.*, 225, p. 1169–1179, 1994.
- [192] Varnai P., Djuranovic D., Lavery R., and Hartmann B. Alpha/gamma transitions in the B-DNA backbone. *Nucleic Acids Res.*, 30(24), p. 5398–5406, 2002.
- [193] Verlet L. Computer experiments on classical fluids. i thermodynamical properties of lennard-jones molecules. *Phys. Chem.*, 159, p. 98, 1967.
- [194] Voet D. and Voet J.G. *Biochemistry. Second edition*. John Wiley & Sons, Inc, 1995.
- [195] Wang A., Hakokshima T., Marel G., Boom J., and Rich A. AT base pairs are less stable than GC base pairs in Z-DNA : the crystal structure of d(m⁵CGTA m⁵CG). *Cell*, 37, p. 321–331, 1984.
- [196] Wang J., Cieplak P., and Kollman P. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules ? *J. Comput. Chem.*, 21(12), p. 1049–1074, 2000.
- [197] Wang J., Wang W., and Kollman P. AntechAMBER : An accessory software package for molecular mechanical calculations. *Abstracts of Papers of the American Chemical Society*, 222, 2001.

- [198] Wang Y., Thomas G., and Peticolas W. Sequence dependent conformations of oligomeric DNA's in aqueous solutions and in crystals. *J. Biomol. Struct. Dyn.*, 5, p. 249–274, 1987.
- [199] Watson J. and Crick F. Molecular structure of nucleic acids : a structure for deoxyribose nucleic acid. *Nature*, 171, p. 737, 1953.
- [200] Weiner Paul K. and Kollman Peter A. AMBER : Assisted Model Building with Energy Refinement. A general program for modeling molecules and their interactions. *J. Comput. Chem*, 2(3), p. 287–303, 1981.
- [201] Weiner S.J., Kollman P.A., Case D.A., Singh U.C., Ghio C., Alagona G., Profeta S., and Weiner P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.*, 106(3), p. 765–784, 1984.
- [202] Westhof E. and Sundaralingam M. . A method for the analysis of puckering disorder in five membered rings : the relative mobilities of furanose and proline rings and their effects on polynucleotide and polypeptide backbone flexibility. *J. Am. Chem. Soc.*, 105, p. 970–976, 1983.
- [203] Whitehead J. and Lippard S. Proteins that bind to and mediate the biological activity of platinum anticancer drug-DNA adducts. *Met. Ions Biol. Syst.*, 32, p. 687–726, 1996.
- [204] Wijmenga S. S., Mooren M. M. W., and Hilbers C. W. In Roberts G. C. K., editor, *NMR of Macromolecules : A Practical Approach*. Oxford University Press, New. York, 1993.
- [205] Wijmenga S. S. and Buuren B. N. M. The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Mag. Res. Spect.*, 32(4), p. 287–387, June 1998.
- [206] Wilkins M., Stokes A., and Wilson H. Molecular structure of the nucleic acids : molecular structure of deoxypentose nucleic acids. *Nature*, 171, p. 738–740, 1953.
- [207] Wong E. and Giandomenico C.M. Current status of platinum-based antitumor drugs. *Chem. Rev.*, 99(9), p. 2451–66, 1999.
- [208] Wu L., Schwederski B.E., and Margerum D.W. Stepwise hydrolysis kinetics of tetrachloroplatinate(II) in base. *Inorg. Chem.*, 29, p. 3578–3584, 1990.
- [209] Yaneva J., Leuba S.H., Holde K., and Zlatanova J. The major chromatin protein histone H1 binds preferentially to cis-platinum-damaged DNA. *Proc. Natl. Acad. Sci. USA*, 94(25), p. 13448–13451, 1997.

- [210] Yang D., Boom S. S. G. E., Reedijk J., Boom J. H., and Wang A. H.-J. Structure and isomerization of an intrastrand cisplatin-cross-linked octamer DNA duplex by NMR analysis. *Biochemistry*, 34, p. 12912–12920, 1995.
- [211] Yao S., Plastaras J. P., and Marzilli L. G. A molecular mechanics AMBER-type force field for modeling platinum complexes of guanine derivatives. *Inorg. Chem.*, 33, p. 6061–6077, 1994.
- [212] Yarema K., Lippard S., and Essigmann J. Mutagenic and genotoxic effects of DNA adducts formed by the anticancer drug cis-diamminedichloroplatinum(II). *Nucleic Acids Res.*, 23(20), p. 4066–4072, 1995.
- [213] Zamble D. and Lippard S. Cisplatin and DNA repair in cancer chemotherapy. *Trends Biochem. Sci.*, 20(10), p. 435–439, October 1995.
- [214] Zhen W., Link C.J., O'Connor P.M., Reed E., Parker R., Howell S.B., and Bohr V.A. Increased gene-specific repair of cisplatin interstrand cross-links in cisplatin-resistant human ovarian cancer cell lines. *Mol. Cell. Biol.*, 12(9), p. 3689–98, 1992.
- [215] Zhou N., Manogaran S., Zon G., and James T. Deoxyribose ring conformation of d(GGTATACC)₂ : an analysis of vicinal proton-proton coupling constants from two-dimensional proton nuclear magnetic resonance. *Biochemistry*, 27(16), p. 6013–6020, 9 August 1988.
- [216] Zlatanova J., Yaneva J., and Leuba S.H. Proteins that specifically recognize cisplatin-damaged DNA : a clue to anticancer activity of cisplatin. *FASEB J.*, 12(10), p. 791–799, July 1998.

Annexes

Modifications pour mettre en place la nouvelle paramétrisation du cisplatine

12a13,14

> CP 12.01 ***ST atome du cycle furanose pour deoxy

> CS 12.01 ***ST atome du cycle furanose pour deoxy

64c66

< C H HO N NA NB NC N2 NT N2 N3 NE NF NS NW N* O OH OS P O2

—

> C H HO N NA NB NC N2 NT N2 N3 NE NF NS NW N* O OH OS P O2

158a161,169

> CP-CP 310.0 1.526 ***ST comme JCC,7,(1986),230; AA, SUGARS

> CP-CS 310.0 1.526 ***ST comme JCC,7,(1986),230; AA, SUGARS

> CT-CP 310.0 1.526 ***ST comme JCC,7,(1986),230; AA, SUGARS

> CP-HC 340.0 1.090 ***ST comme changed from 331 bsd on NMA nmodes; AA, SUGARS

> CS-HC 340.0 1.090 ***ST comme changed from 331 bsd on NMA nmodes; AA, SUGARS

> CP-H1 340.0 1.090 ***ST comme changed from 331 bsd on NMA nmodes; AA, RIBOSE

> CP-H2 340.0 1.090 ***ST comme changed from 331 bsd on NMA nmodes; SUGARS

> CP-N* 337.0 1.475 ***ST comme JCC,7,(1986),230; ADE,CYT,GUA,THY,URA

> CP-OS 320.0 1.410 ***ST comme JCC,7,(1986),230; NUCLEIC ACIDS

376a388,417

> CT-CP-CP 40.0 109.50 ***ST comme CT-CT-CT

> CP-CP-CP 40.0 109.50 ***ST comme CT-CT-CT

> CP-CP-CS 40.0 109.50 ***ST comme CT-CT-CT

> CP-CS-CP 40.0 109.50 ***ST comme CT-CT-CT

> CP-CP-HC 50.0 109.50 ***ST comme changed based on NMA nmodes

> CP-CS-HC 50.0 109.50 ***ST comme changed based on NMA nmodes

> CP-CP-H1 50.0 109.50 ***ST comme changed based on NMA nmodes

> CT-CP-H1 50.0 109.50 ***ST comme changed based on NMA nmodes

> CP-CT-H1 50.0 109.50 ***ST comme changed based on NMA nmodes

> CT-CP-H1 50.0 109.50 ***ST comme changed based on NMA nmodes

> CS-CP-H1 50.0 109.50 ***ST comme changed based on NMA nmodes

> CP-CP-H2 50.0 109.50 ***ST comme changed based on NMA nmodes

> CS-CP-H2 50.0 109.50 ***ST comme changed based on NMA nmodes

> CP-CP-N* 50.0 109.50 ***ST comme CT-CT-N*

> CS-CP-N* 50.0 109.50 ***ST comme CT-CT-N*

> CP-CP-OS 50.0 109.50 ***ST comme CT-CT-OS

> CP-CT-OS 50.0 109.50 ***ST comme CT-CT-OS

> CT-CP-OS 50.0 109.50 ***ST comme CT-CT-OS

> CS-CP-OS 50.0 109.50 ***ST comme CT-CT-OS

> HC-CP-HC 35.0 109.50 ***ST comme HC-CT-HC

> HC-CS-HC 35.0 109.50 ***ST comme HC-CT-HC

> H2-CP-N* 50.0 109.50 ***ST comme changed based on NMA nmodes

> H1-CP-OS 50.0 109.50 ***ST comme changed based on NMA nmodes

> H2-CP-OS 50.0 109.50 ***ST comme changed based on NMA nmodes

```

> N*-CP-OS 50.0 109.50 ***ST comme N*-CT-OS
> CB-N*-CP 70.0 125.80 ***ST comme CB-N*-CT
> CK-N*-CP 70.0 128.80 ***ST comme CK-N*-CT
> CM-N*-CP 70.0 121.20 ***ST comme CM-N*-CT
> CP-OS-CP 60.0 109.50 ***ST comme CT-OS-CT
> CP-OS-P 100.0 120.50 ***ST comme CT-OS-P
437c478
< CT-OS-CT-N* 1 1.0 000.0 2.
—
> CT-OS-CT-N* 1 0.65 000.0 2. ***ST comme parm 99 avant :
452c493
< OS-CT-CT-OS 1 1.50 0.0 2. pucker anal (93')
—
> OS-CT-CT-OS 1 1.175 0.0 2. ***ST comme parm 99 avant : pucker anal (93')
454c495
< OS-CT-CT-OH 1 1.50 0.0 2. pucker anal (93')
—
> OS-CT-CT-OH 1 1.175 0.0 2. ***ST comme parm 99 avant : pucker anal (93')
456c497
< OH-CT-CT-OH 1 1.50 0.0 2. check glycolWC? puc
—
> OH-CT-CT-OH 1 1.175 0.0 2. ***ST comme parm 99 avant : check glycolWC? puc
482a524,544
> X -CP-CP-X 9 1.40 0.0 3. ***ST comme JCC,7,(1986),230
> X -CT-CP-X 9 1.40 0.0 3. ***ST comme JCC,7,(1986),230
> X -CS-CP-X 9 1.40 0.0 3. ***ST comme JCC,7,(1986),230
> X -CP-N*-X 6 0.00 0.0 2. ***ST comme JCC,7,(1986),230
> X -CP-OS-X 3 1.15 0.0 3. ***ST comme JCC,7,(1986),230
> CP-CP-OS-CP 1 0.383 0.0 -3. ***ST comme CT-CT-OS-CT
> CS-CP-OS-CP 1 0.383 0.0 -3. ***ST comme CT-CT-OS-CT
> CP-CP-OS-CP 1 0.1 180.0 2. ***ST comme CT-CT-OS-CT
> CS-CP-OS-CP 1 0.1 180.0 2. ***ST comme CT-CT-OS-CT
> CP-OS-CP-N* 1 0.383 0.0 -3. ***ST comme CT-OS-CT-N*
> CP-OS-CP-N* 1 0.65 000.0 2. ***ST comme CT-OS-CT-N* (parm 99)
> CS-CP-OS-P 3 1.15 0.0 -3. ***ST comme JCC,7,(1986),230
> CS-CP-OS-P 0 0.00 -60.0 2. ***ST paramétrisation 29.09.2003
> OS-CP-CP-OS 1 0.144 0.0 -3. ***ST comme JCC,7,(1986),230
> OS-CP-CP-OS 1 1.175 0.0 2. ***ST comme pucker anal (93')
> OS-CS-CT-OS 1 0.144 0.0 -3. ***ST comme JCC,7,(1986),230
> OS-CS-CT-OS 1 1.175 0.0 2. ***ST comme pucker anal (93')
> OS-CP-N*-CK 1 0.00 000.0 -2. ***ST comme sugar frag calc (PC)
> OS-CP-N*-CK 1 2.50 0.0 1. ***ST comme sugar frag calc (PC)
> OS-CP-N*-CM 1 0.00 000.0 -2. ***ST comme sugar frag calc (PC)
> OS-CP-N*-CM 1 2.50 0.0 1. ***ST comme sugar frag calc (PC)
501a564
> CK-CB-N*-CP 1.0 180. 2. ***ST comme CK-CB-N*-CT

```

519c582

< C C* CA CB CC CN CM CK CQ CW CV CR CA CX CY CD

—

> C C* CA CB CC CN CM CK CQ CW CV CR CA CX CY CD CP CS

540a604,605

> CP 1.9080 0.1094 ***ST comme Spellmeyer

> CS 1.9080 0.1094 ***ST comme Spellmeyer

Fichier d'entrée pour le module PREP

0 0 2

D-GUANOSINE - with 5' - phosphate group and 3' - O(minus) group

GDA

GDA INT 1

CORRECT NOMIT DU BEG

0.0

1 DUMM DU M 0 -1 -2 0.00 0.00 0.00 0.0000

2 DUMM DU M 1 0 -1 1.00 0.00 0.00 0.0000

3 DUMM DU M 2 1 0 1.00 90.00 0.00 0.0000

4 P P M 3 2 1 1.60 119.04 200.00 1.1659

5 O1P O2 E 4 3 2 1.48 109.61 150.00 -0.7761

6 O2P O2 E 4 3 2 1.48 109.58 20.00 -0.7761

7 O5' OS M 4 3 2 1.60 101.43 -98.89 -0.4954

8 C5' CT M 7 4 3 1.44 119.00 -39.22 -0.0069

9 H5'1 H1 E 8 7 4 1.09 109.50 60.00 0.0754

10 H5'2 H1 E 8 7 4 1.09 109.50 -60.00 0.0754

11 C4' CP M 8 7 4 1.52 110.00 180.00 0.1629

12 H4' H1 E 11 8 7 1.09 109.50 -200.00 0.1176

13 O4' OS S 11 8 7 1.46 108.86 -86.31 -0.3691

12 C1' CP B 11 9 5 1.42 110.04 105.60 0.077265

13 H1' H2 E 12 11 9 1.09 109.50 -240.00 0.216065

14 N9 N* S 12 11 9 1.49 108.06 -127.70 0.162994

15 C8 CK B 14 12 11 1.38 129.20 81.59 -0.175248

16 H8 H5 E 15 14 12 1.08 120.00 0.00 0.189499

17 N7 NE S 15 14 12 1.31 114.00 -179.90 0.018089

18 C5 CB S 17 15 14 1.39 103.90 0.00 -0.211378

19 C6 C B 18 17 15 1.42 130.40 180.00 0.595141

20 O6 O E 19 18 17 1.23 128.80 0.00 -0.466742

21 N1 NA B 19 18 17 1.40 111.38 180.00 -0.669066

22 H1 H E 21 19 18 1.00 117.36 179.90 0.400762

23 C2 CA B 21 19 18 1.38 125.24 -0.10 0.929895

24 N2 N2 B 23 21 19 1.34 116.02 180.00 -0.969660

25 H21 H E 24 23 21 1.01 127.00 -0.82 0.459672

26 H22 H E 24 23 21 1.01 116.53 -179.44 0.459672

27 N3 NC S 23 21 19 1.33 123.30 0.00 -0.668746


```

28 C4 CB E 27 23 21 1.36 112.20 0.00 0.348341
31 C3' CP M 11 8 7 1.53 115.78 -329.11 0.0713
32 H3' H1 E 31 11 8 1.09 109.50 30.00 0.0985
33 C2' CS B 31 11 8 1.53 102.80 -86.30 -0.0854
34 H2'1 HC E 33 31 11 1.09 109.50 120.00 0.0718
35 H2'2 HC E 33 31 11 1.09 109.50 240.00 0.0718
36 O3' OS M 31 11 8 1.42 116.52 -203.47 -0.5232
IMPROPER
C8 C4 N9 C1'
C5 N1 C6 O6
C6 C2 N1 H1
C2 H21 N2 H22
N7 N9 C8 H8
N1 N3 C2 N2
LOOP CLOSING EXPLICIT
C1' C2'
C4 C5
C4 N9
DONE
D-GUANOSINE - with 5' - phosphate group and 3' - OH group. Charges GG_Pt/DFT 18-Feb-97
GDB
GDB INT 1
CORRECT OMIT DU BEG
0.0
1 DUMM DU M 0 -1 -2 0.00 0.00 0.00 0.0000
2 DUMM DU M 1 0 -1 1.00 0.00 0.00 0.0000
3 DUMM DU M 2 1 0 1.00 90.00 0.00 0.0000
4 P P M 3 2 1 1.60 119.04 200.00 1.1659
5 O1P O2 E 4 3 2 1.48 109.61 150.00 -0.7761
6 O2P O2 E 4 3 2 1.48 109.58 20.00 -0.7761
7 O5' OS M 4 3 2 1.60 101.43 -98.89 -0.4954
8 C5' CT M 7 4 3 1.44 119.00 -39.22 -0.0069
9 H5'1 H1 E 8 7 4 1.09 109.50 60.00 0.0754
10 H5'2 H1 E 8 7 4 1.09 109.50 -60.00 0.0754
11 C4' CP M 8 7 4 1.52 110.00 180.00 0.1629
12 H4' H1 E 11 8 7 1.09 109.50 -200.00 0.1176
13 O4' OS S 11 8 7 1.46 108.86 -86.31 -0.3691
12 C1' CP B 11 9 5 1.42 110.04 105.60 0.077265
13 H1' H2 E 12 11 9 1.09 109.50 -240.00 0.216065
14 N9 N* S 12 11 9 1.49 108.06 -127.70 0.162994
15 C8 CK B 14 12 11 1.38 129.20 81.59 -0.175248
16 H8 H5 E 15 14 12 1.08 120.00 0.00 0.189499
17 N7 NF S 15 14 12 1.31 114.00 -179.90 0.018089
18 C5 CB S 17 15 14 1.39 103.90 0.00 -0.211378
19 C6 C B 18 17 15 1.42 130.40 180.00 0.595141
20 O6 O E 19 18 17 1.23 128.80 0.00 -0.466742

```

```

21 N1 NA B 19 18 17 1.40 111.38 180.00 -0.669066
22 H1 H E 21 19 18 1.00 117.36 179.90 0.400762
23 C2 CA B 21 19 18 1.38 125.24 -0.10 0.929895
24 N2 N2 B 23 21 19 1.34 116.02 180.00 -0.969660
25 H21 H E 24 23 21 1.01 127.00 -0.82 0.459672
26 H22 H E 24 23 21 1.01 116.53 -179.44 0.459672
27 N3 NC S 23 21 19 1.33 123.30 0.00 -0.668746
28 C4 CB E 27 23 21 1.36 112.20 0.00 0.348341
31 C3' CP M 11 8 7 1.53 115.78 -329.11 0.0713
32 H3' H1 E 31 11 8 1.09 109.50 30.00 0.0985
33 C2' CS B 31 11 8 1.53 102.80 -86.30 -0.0854
34 H2'1 HC E 33 31 11 1.09 109.50 120.00 0.0718
35 H2'2 HC E 33 31 11 1.09 109.50 240.00 0.0718
36 O3' OS M 31 11 8 1.42 116.52 -203.47 -0.5232
IMPROPER
C8 C4 N9 C1'
C5 N1 C6 O6
C6 C2 N1 H1
C2 H21 N2 H22
N7 N9 C8 H8
N1 N3 C2 N2
LOOP CLOSING EXPLICIT
C1' C2'
C4 C5
C4 N9
DONE
PTJ
PTJ
PTJ INT 1
CORRECT OMIT DU BEG
0.0
1 DUMM DU M 0 0 0 0. 0. 0. 0. charges updated for GUA, GUA et Pt 18.02.1997
2 DUMM DU M 0 0 0 1. 0. 0. 0. according to DFT 6-31g** calculations
3 DUMM DU M 0 0 0 1. 90. 0. 0. H3=0.346766 + H4=0.388340 / 2 = 0.367553
4 PT PT M 0 0 0 2.00 128.1 180. 0.047440 Mis-à-jour par Stef le 17/12/2003 pour N3P
5 N3P NS 3 0 0 0 2.04 90. 90. -0.701331 et N4P :
6 H31 HT E 0 0 0 1.00 109. 180. 0.367553 N3P=-0.635233 + N4P=-0.767429 / 2 = -0.701331
7 H32 HT E 0 0 0 1.00 109. 60. 0.367553
8 H33 HT E 0 0 0 1.00 109. -60. 0.367553
9 N4P NW 3 0 0 0 2.04 180. 180. -0.701331
10 H41 HT E 0 0 0 1.00 109. 180. 0.367553
11 H42 HT E 0 0 0 1.00 109. 60. 0.367553
12 H43 HT E 0 0 0 1.00 109. -60. 0.367553
DONE
STOP

```

Fichier d'entrée pour le module LINK

```

Titre du fichier base.linkin
GDA 1GDA GDB 1GDB PTJ 1PTJ
DU 0 0 0 0 DNA(5'-3') D 0 0 1 3 0 DG5 2DC 2DC 2GDA 2GDB 2DG 2DT 2DC 2DG 2DC3 2
DNA(5'-3') D 0 0 1 3 0 DG5 2DC 2DG 2DA 2DC 2DC 2DC 2DG 2DG 2DC3 2
PTJ(Pt(NH3)2)2+ O 1 0 1 630 PTJ 2
4 1N7 PT 1 5 1N7 PT 1
QUIT

```

Fichier d'entrée pour le module LINK_ions

```

Titre du fichier base.linkin
GDA 1GDA GDB 1GDB PTJ 1PTJ
DU 0 0 0 0 DNA(5'-3') D 0 0 1 3 0 DG5 2DC 2DC 2GDA 2GDB 2DG 2DT 2DC 2DG 2DC3 2
DNA(5'-3') D 0 0 1 3 0 DG5 2DC 2DG 2DA 2DC 2DC 2DC 2DG 2DG 2DC3 2
PTJ(Pt(NH3)2)2+ O 1 0 1 630 PTJ 2
4 1N7 PT 1 5 1N7 PT 1
QUIT

```

Fichier d'entrée pour le module EDIT

```

DNA sequence
0 0 0 0
XYZ
OMIT
XRAY
0 0 0 0
XRAY
2 0 0 0
QUIT

```

Fichier d'entrée pour le module EDIT avec l'insertion d'ions

```

DNA sequence
0 0 0 0
XYZ
OMIT
XRAY
0 0 0 0
XRAY
2 0 0 0

```

CION
 1.000 3.000 2.000 2.000
 DA P 1.00 6.00
 DG P 1.00 6.00
 DGDA P 1.00 6.00
 DGDB P 1.00 6.00
 DC P 1.00 6.00
 DT P 1.00 6.00
 DA3 P 1.00 6.00
 DA5 P 1.00 6.00
 DG3 P 1.00 6.00
 DG5 P 1.00 6.00
 DC3 P 1.00 6.00
 DC5 P 1.00 6.00
 DT3 P 1.00 6.00
 DT5 P 1.00 6.00
 A P 1.00 6.00
 G P 1.00 6.00
 C P 1.00 6.00
 T P 1.00 6.00
 A3 P 1.00 6.00
 A5 P 1.00 6.00
 G3 P 1.00 6.00
 G5 P 1.00 6.00
 C3 P 1.00 6.00
 C5 P 1.00 6.00
 T3 P 1.00 6.00
 T5 P 1.00 6.00
 QUIT

Fichier d'entrée pour le module EDIT avec l'ajout d'une boite d'eau

DNA sequence
 0 0 0 0
 XYZ
 OMIT
 XRAY
 0 0 0 0
 XRAY
 2 0 0 0
 BOX
 HW OW 6
 0.41700 2.80000 2.30000

```
10.1000 11.7000 10.8000
QUIT
```

Fichier d'entrée pour le module PARM

```
Titre du fichier .parmin
BIN FOR MOD4CONS
1 1 1
0 0 0
21 622 0 0 0.00 10.00
19 620 0 0 0.00 10.00
17 617 0 0 0.00 10.00
337 309 0 0 0.00 10.00
335 307 0 0 0.00 10.00
333 304 0 0 0.00 10.00
630 108 113 107 180.00 2.50
630 108 113 105 180.00 2.50
630 105 108 118 180.00 15.90
630 105 109 108 180.00 14.98
630 141 146 140 180.00 2.50
630 141 146 138 180.00 2.50
630 138 141 151 180.00 15.90
630 138 142 141 180.00 14.98
```

amber4__1

```
#!/bin/bash
echo '***** prepA4__1 *****' $1
[ -e PTJ ] && rm PTJ
[ -e GDA ] && rm GDA
[ -e GDB ] && rm GDB
[ -e $1.top1 ] && rm $1.top1
[ -e 1.crd1 ] && rm $1.crd1
[ -e $1.prepout1 ] && rm $1.prepout1
prepA4__1 $1
echo '***** linkA4__1 *****' $1
[ -e $1.linkout1 ] && rm $1.linkout1
[ -e $1.linktop1 ] && rm $1.linktop1
linkA4__1 $1
echo '***** editA4__1 *****' $1
[ -e $1.editout1 ] && rm $1.editout1
[ -e $1.editop1 ] && rm $1.editop1
editA4__1 $1
echo '***** parmA4__1 *****' $1
```

```
[ -e $1.parmout1 ] && rm $1.parmout1
parmA4_1 $1
exit
```

prepA4_1

```
prep4.1 -O -i $1.prepin1\
-o $1.prepout1
rm params
exit
```

linkA4_1

```
link4.1 -O -i $1.linkin1\
-o $1.linkout1\
-l $1.linktop1\
-p FF/db94.dat
```

editA4_1

```
edit4.1 -O -i $1.editin1\
-o $1.editout1\
-l $1.linktop1\
-e $1.editop1\
-pi $1.pdbin1 \
-b FF/db94.dat
mv pdbout $1.pdbout
```

parmA4_1

```
parm4.1 -O -i $1.parmin1\
-o $1.parmout1\
-e $1.editop1\
-c $1.crd1\
-p $1.top1\
-f FF/parm99_cisPt_ST.dat
```

amberA4_2

```
#!/bin/bash
echo '***** prepA4_1 *****' $1
[ -e PTJ ] && rm PTJ
```

```

[ -e GDA ] && rm GDA
[ -e GDB ] && rm GDB
[ -e $1.top2 ] && rm $1.top2
[ -e 1.crd2 ] && rm $1.crd2
[ -e $1.prepout2 ] && rm $1.prepout2
prepA4_2 $1
echo '***** linkA4_1 *****' $1
[ -e $1.linkout2 ] && rm $1.linkout2
[ -e $1.linktop2 ] && rm $1.linktop2
linkA4_2 $1
echo '***** editA4_1 *****' $1
[ -e $1.editout2 ] && rm $1.editout2
[ -e $1.editop2 ] && rm $1.editop2
editA4_2 $1
echo '***** parmA4_1 *****' $1
[ -e $1.parmout2 ] && rm $1.parmout2
parmA4_2 $1
exit

```

prepA4_2

```

prep4.1 -O -i $1.prepin2\
-o $1.prepout2
rm params
exit

```

linkA4_2

```

link4.1 -O -i $1.linkin2\
-o $1.linkout2\
-l $1.linktop2\
-p FF/db94.dat

```

editA4_2

```

edit -O -i $1.editin2\
-o $1.editout2\
-l $1.linktop2\
-e $1.editop2\
-pi $1.pdbin2 \
-b ./FF/wat216.dat
mv pdbout $1.pdbout2

```

parmA4_2

```
parm4.1 -O -i $1.parmin2\
-o $1.parmout2\
-e $1.editop2\
-c $1.crd2\
-p $1.top2\
-f FF/parm99_cisPt_ST.dat
```

Script pour lancer les minimisations

```
#!/bin/csh
echo $1
foreach i (1 2 3 4 5 6)
date
echo ' min. #' $i
sander_classic -O \
-i 'mincomH2O_'$i.dat \
-p $1.top2 \
-c $1.rst \
-ref $1.rst \
-o $1.$i'min_outH2O' \
-r $1.'min'$i
\cat mdinfo >> $1.bilan_6min_H2O
\cp $1.'min'$i $1.rst
end
```

mincomH2O_1.dat

```
GGG-Pt1,2
&cntrl
IMIN=1, TIMLIM=999999, NSNB=50, IG=71277, IBELLY=0,
NTX=1, IREST=0, NTF=1, NTC=1, NTN=1, NTID=0,
IDIEL=1, DIELC=1, CUT=9.0, SCNB=2.0, SCOE= 1.2, NTB=1, IEWALD=1,
MAXCYC=500, NCYC=10, NTMIN=1, DRMS=0.000001,
NMRMAX=0, NTPR=10,
NTR=1,
&end
65.2688 53.6024 53.1546 90.0 90.0 90.0
64 54 54 4 0 0 0
0.000001
Contraintes sur tout l'ADN + Pt
100.0
RES 1 21
```



```

END
Contraintes sur tous les ions Na+
100.0
RES 22 37
END
END

```

mincomH2O_2.dat

```

GGG-Pt1,2
&cntrl
IMIN=1, TIMLIM=999999, NSNB=50, IG=45781, IBELLY=0,
NTX=1, IREST=0, NTF=1, NTC=1, NTNB=1, NTID=0,
IDIEL=1, DIELC=1, CUT=9.0, SCNB=2.0, SC EE= 1.2, NTB=1, IEWALD=1,
MAXCYC=500, NCYC=10, NTMIN=1, DRMS=0.000001,
NMRMAX=0, NTPR=10,
NTR=1,
&end
65.2688 53.6024 53.1546 90.0 90.0 90.0
64 54 54 4 0 0 0
0.000001
Contraintes sur tout l'ADN + Pt
50.0
RES 1 21
END
Contraintes sur tous les ions Na+
50.0
RES 22 37
END
END

```

mincomH2O_3.dat

```

GGG-Pt1,2
&cntrl
IMIN=1, TIMLIM=999999, NSNB=50, IG=78365, IBELLY=0,
NTX=1, IREST=0, NTF=1, NTC=1, NTNB=1, NTID=0,
IDIEL=1, DIELC=1, CUT=9.0, SCNB=2.0, SC EE= 1.2, NTB=1, IEWALD=1,
MAXCYC=250, NCYC=10, NTMIN=1, DRMS=0.000001,
NMRMAX=0, NTPR=10,
NTR=1,
&end
65.2688 53.6024 53.1546 90.0 90.0 90.0
64 54 54 4 0 0 0

```

```

0.000001
Contraintes sur tout l'ADN + Pt
25.0
RES 1 21
END
Contraintes sur tous les ions Na+
25.0
RES 22 37
END
END

```

mincomH2O__4.dat

```

GGG-Pt1,2
&cntrl
IMIN=1, TIMLIM=999999, NSNB=50, IG=14598, IBELLY=0,
NTX=1, IREST=0, NTF=1, NTC=1, NTNB=1, NTID=0,
IDIEL=1, DIELC=1, CUT=9.0, SCNB=2.0, SC EE= 1.2, NTB=1, IEWALD=1,
MAXCYC=250, NCYC=10, NTMIN=1, DRMS=0.000001,
NMRMAX=0, NTPR=10,
NTR=1,
&end
65.2688 53.6024 53.1546 90.0 90.0 90.0
64 54 54 4 0 0 0
0.000001
Contraintes sur tout l'ADN + Pt
10.0
RES 1 21
END
Contraintes sur tous les ions Na+
10.0
RES 22 37
END
END

```

mincomH2O__5.dat

```

GGG-Pt1,2
&cntrl
IMIN=1, TIMLIM=999999, NSNB=50, IG=24513, IBELLY=0,
NTX=1, IREST=0, NTF=1, NTC=1, NTNB=1, NTID=0,
IDIEL=1, DIELC=1, CUT=9.0, SCNB=2.0, SC EE= 1.2, NTB=1, IEWALD=1,
MAXCYC=250, NCYC=10, NTMIN=1, DRMS=0.000001,
NMRMAX=0, NTPR=10,

```

```

NTR=1,
&end
65.2688 53.6024 53.1546 90.0 90.0 90.0
64 54 54 4 0 0 0
0.000001
Contraintes sur tout l'ADN + Pt
5.0
RES 1 21
END
END

```

mincomH2O__6.dat

```

GGG-Pt1,2
&cntrl
IMIN=1, TIMLIM=999999, NSNB=50, IG=65127, IBELLY=0,
NTX=1, IREST=0, NTF=1, NTC=1, NTNB=1, NTID=0,
IDIEL=1, DIELC=1, CUT=9.0, SCNB=2.0, SC EE= 1.2, NTB=1, IEWALD=1,
MAXCYC=500, NCYC=10, NTMIN=1, DRMS=0.000001,
NMRMAX=0, NTPR=10,
NTR=0,
&end
65.2688 53.6024 53.1546 90.0 90.0 90.0
64 54 54 4 0 0 0
0.000001

```

Lancement de la dynamique

```

#!/bin/tcsh
if(($#argv != 2)) then
echo "Le nombre d'arguments est diff. a deux. ERREUR \!\!"
exit (1)
endif
\rm $1/*.file
\rm $1/*.rst
set DO_PARALLEL='/usr/local/mpich-1.2.2.3/bin/mpirun -np 6 -nolocal -machinefile mach'
#
# Fichiers et répertoires de sortie
#
set crd = $1/crd
if (!( -e $crd )) then
mkdir $crd
endif
set lst = $1/lst

```

```
if (!(-e $lst)) then
mkdir $lst
endif
set rst = $1/rst
if !(-e $rst)) then
mkdir $rst
endif
set out = $1/$2'_run'
if (-e $out) then
\rm $out
endif
echo "Arguments : $1 - $2" > $out
echo "Début de la dynamique : \t\t\t\t\t't'date'" >> $out
set dirbas = ./DM_Tcte300_H2O
\cp $1/prm/$2.crd $1/$2.rst
#
# heating : etape #1
#
foreach i ( 1 2 )
set tag = he300K$i
set file = $2.$tag
echo "Heating #$i\t\t\t\t\t$file\t\t't'date'" >> $out
$DO_PARALLEL /usr/local/Programmes/Ambber6/amber6/exe/sander -O \
-i $dirbas/$tag \
-p $1/prm/$2.top \
-c $1/$2.rst \
-ref $1/$2.rst \
-o $1/lst/'$file'_l' \
-x $1/crd/'$file'_c' \
-r $1/$2.rst_he$i
\cp $1/$2.rst_he$i $1/$2.rst
\mv $1/$2.rst_he$i $1/rst
if (-e $1/rst/$2.rst_he$i.Z) \rm $1/rst/$2.rst_he$i.Z
compress $1/rst/$2.rst_he$i
if (-e $1/crd/'$file'_c'.Z) \rm $1/crd/'$file'_c'.Z
compress $1/crd/'$file'_c'
if($status) then
echo "Arrêt inopiné de la dynamique à l'étape : $tag" >> $out
exit(1)
endif
end
#
# equilibration : etape #2
#
foreach i (1 2 3 4 5 6 7)
set tag = ei300K$i
```



```

if (-e $1/rst/$2.rst_ri1.Z) \rm $1/rst/$2.rst_ri1.Z
compress $1/rst/$2.rst_ri1
if (-e $1/crd/$file'_c'.Z) \rm $1/crd/$file'_c'.Z
compress $1/crd/$file'_c'
if($status) then
echo "Arrêt inopiné de la dynamique à l'étape : $tag" >> $out
exit(1)
endif
#
# equilibration : etape #4
#
foreach i (8)
set tag = ei300K$i
set file = $2.$tag
echo "Equilibration # $i \t \t $file \t \t 'date'" >> $out
$DO_PARALLEL /usr/local/Programmes/Amber6/amber6/exe/sander -O \
-i $dirbas/$tag \
-p $1/prm/$2.top \
-c $1/$2.rst \
-ref $1/$2.rst \
-o $1/lst/$file'_l' \
-x $1/crd/$file'_c' \
-r $1/$2.rst_ei$i
\cp $1/$2.rst_ei$i $1/$2.rst
\mv $1/$2.rst_ei$i $1/rst
if (-e $1/rst/$2.rst_ei$i.Z) \rm $1/rst/$2.rst_ei$i.Z
compress $1/rst/$2.rst_ei$i
if (-e $1/crd/$file'_c'.Z) \rm $1/crd/$file'_c'.Z
compress $1/crd/$file'_c'
if($status) then
echo "Arrêt inopiné de la dynamique à l'étape : $tag" >> $out
exit(1)
endif
end
#
# equilibration : etape #5. IG diff, Prog. de Young pour le CM.
#
set i = 1
while ($i < 2)
set tag = prodstb$i
set file = $2.$tag
echo "Equilibration # $i \t \t $file \t \t 'date'" >> $out
$DO_PARALLEL /usr/local/Programmes/Amber6/amber6/exe/sander -O \
-i $dirbas/prodstb \
-p $1/prm/$2.top \
-c $1/$2.rst \

```

```

-ref $1/$2.rst \
-o $1/lst/$file'_l' \
-x $1/crd/$file'_c' \
-r $1/$2.rst__prodstb$i
\cp $1/$2.rst__prodstb$i $1/$2.rst
\mv $1/$2.rst__prodstb$i $1/rst
if (-e $1/rst/$2.rst__prodstb$i.Z) \rm $1/rst/$2.rst__prodstb$i.Z
compress $1/rst/$2.rst__prodstb$i
if (-e $1/crd/$file'_c'.Z) \rm $1/crd/$file'_c'.Z
compress $1/crd/$file'_c'
if($status) then
echo "Arrêt inopiné de la dynamique à l'étape : $tag" >> $out
exit(1)
endif
@ i++
end
#
# Comment récupérer la taille de la boîte sur la fin du rst ...
#
set boxx='tail -1 $2.rst | cut -c 3-12'
set boxy='tail -1 $2.rst | cut -c 15-24'
set boxz='tail -1 $2.rst | cut -c 27-36'
sed 's/a=.*, b=.*, c=.*,$/a='$boxx', b='$boxy', c='$boxz',/g' $dirbas/ri300K_Maxwell2 > toto
mv toto $dirbas/ri300K_Maxwell2
#
# random technique : etape #6. Maxwell-Boltzman distribution (Andersen option NTT=-1 et IG
dif.)
#
set tag = ri300K_Maxwell2
set file = $2.$tag
echo "Maxwell-Boltzman distribution OK \!\!\!\t$file\t'date'" >> $out
$DO_PARALLEL /usr/local/Programmes/Amber6/amber6/exe/sander -O \
-i $dirbas/$tag \
-p $1/prm/$2.top \
-c $1/$2.rst \
-ref $1/$2.rst \
-o $1/lst/$file'_l' \
-x $1/crd/$file'_c' \
-r $1/$2.rst__ri2
\cp $1/$2.rst__ri2 $1/$2.rst
\mv $1/$2.rst__ri2 $1/rst
if (-e $1/rst/$2.rst__ri2.Z) \rm $1/rst/$2.rst__ri2.Z
compress $1/rst/$2.rst__ri2
if (-e $1/crd/$file'_c'.Z) \rm $1/crd/$file'_c'.Z
compress $1/crd/$file'_c'
if($status) then

```

```

echo "Arrêt inopiné de la dynamique à l'étape : $tag" >> $out
exit(1)
endif
#
# equilibration : etape #7. IG diff, Prog. de Young pour le CM.
#
set i = 2
while ($i < 3)
set tag = prodstb$i
set file = $2.$tag
echo "Equilibration # $i \t \t $file \t \t `date`" >> $out
$DO_PARALLEL /usr/local/Programmes/Amber6/amber6/exe/sander -O \
-i $dirbas/prodstb \
-p $1/prm/$2.top \
-c $1/$2.rst \
-ref $1/$2.rst \
-o $1/1st/$file'_l' \
-x $1/crd/$file'_c' \
-r $1/$2.rst_prodstb$i
\cp $1/$2.rst_prodstb$i $1/$2.rst
\mv $1/$2.rst_prodstb$i $1/rst
if (-e $1/rst/$2.rst_prodstb$i.Z) \rm $1/rst/$2.rst_prodstb$i.Z
compress $1/rst/$2.rst_prodstb$i
if (-e $1/crd/$file'_c'.Z) \rm $1/crd/$file'_c'.Z
compress $1/crd/$file'_c'
if($status) then
echo "Arrêt inopiné de la dynamique à l'étape : $tag" >> $out
exit(1)
endif
@ i++
end
#
# Comment récupérer la taille de la boîte sur la fin du rst ...
#
set boxx='tail -1 $2.rst | cut -c 3-12'
set boxy='tail -1 $2.rst | cut -c 15-24'
set boxz='tail -1 $2.rst | cut -c 27-36'
sed 's/a=.*, b=.*, c=.*,/a='$boxx', b='$boxy', c='$boxz',/g' $dirbas/ri300K_Maxwell3 > toto
mv toto $dirbas/ri300K_Maxwell3
#
# random technique : etape #8. Maxwell-Boltzman distribution (Andersen option NTT=-1 et IG
dif.)
#
set tag = ri300K_Maxwell3
set file = $2.$tag
echo "Maxwell-Boltzman distribution OK \!\!\! \t $file \t `date`" >> $out

```



```

TIMLIM=999999, IG=71277, IBELLY=0, NTR=1,
NTB=1,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=2500, DT=0.002, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=150.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=0, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=1,
&end
&ewald,
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum_tol=0.000001, &end
#Chauffage de 0 a 150 K en 1 etape de 5 ps.
&wt TYPE='TEMP0', ISTEP1=0, ISTEP2=2500, VALUE1=150.0, VALUE2=300.0, &end
&wt TYPE='END', &end
&rst iat=0, &end
Contraintes sur tout l'ADN
25.0
RES 1 21
END
Contraintes sur tous les ions Na+
25.0
RES 22 37
END
END

```

ei300K1

```

DNA
&cntrl
IMIN=0, NTX=5, IREST=1,
TIMLIM=999999, IG=51277, IBELLY=0, NTR=1,
NTB=1,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=2500, DT=0.002, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=0, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=1,
&end
&ewald,
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum_tol=0.000001, &end

```

```

&wt TYPE='END', &end
&rst iat=0, &end
Contraintes sur tout l'ADN
25.0
RES 1 21
END
Contraintes sur tous les ions Na+
25.0
RES 22 37
END
END

```

ei300K2

```

DNA
&cntrl
IMIN=0, NTX=5, IREST=1,
TIMLIM=999999, IG=89472, IBELLY=0, NTR=1,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=1250, DT=0.002, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=50,
NMROPT=1,
&end
&ewald,
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum__tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end
Contraintes sur tout l'ADN
25.0
RES 1 21
END
Contraintes sur tous les ions Na+
25.0
RES 22 37
END
END

```

ei300K3

```

DNA
&cntrl
IMIN=0, NTX=7, IREST=1,
TIMLIM=999999, IG=26541, IBELLY=0, NTR=1,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=1250, DT=0.002, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=1,
&end
&ewald,
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum_tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end
Contraintes sur tout l'ADN
20.0
RES 1 21
END
Contraintes sur tous les ions Na+
15.0
RES 22 37
END
END

```

ei300K4

```

DNA
&cntrl
IMIN=0, NTX=7, IREST=1,
TIMLIM=999999, IG=11254, IBELLY=0, NTR=1,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=1250, DT=0.002, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=1,
&end

```

```

&ewald,
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54,order=4, dsum__tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end
Contraintes sur tout l'ADN
15.0
RES 1 21
END
Contraintes sur tous les ions Na+
10.0
RES 22 37
END
END

```

ei300K5

```

DNA
&cntrl
IMIN=0, NTX=7, IREST=1,
TIMLIM=999999, IG=66429, IBELLY=0, NTR=1,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=1250, DT=0.002, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=1,
&end
&ewald,
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54,order=4, dsum__tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end
Contraintes sur tout l'ADN
10.0
RES 1 21
END
Contraintes sur tous les ions Na+
5.0
RES 22 37
END
END

```

ei300K6

```

DNA
&cntrl
IMIN=0, NTX=7, IREST=1,
TIMLIM=999999, IG=66429, IBELLY=0, NTR=1,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=1250, DT=0.002, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=1,
&end
&ewald,
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum_tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end
Contraintes sur tout l'ADN
5.0
RES 1 21
END
Contraintes sur tous les ions Na+
2.5
RES 22 37
END
END

```

ei300K7

```

DNA
&cntrl
IMIN=0, NTX=7, IREST=1,
TIMLIM=999999, IG=33447, IBELLY=0, NTR=1,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=5000, DT=0.002, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=1,
&end

```

```

&ewald,
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum_tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end
Contraintes sur tout l'ADN
2.5
RES 1 21
END
END

```

ri300K_Maxwell1

```

DNA
&cntrl
IMIN=0, NTX=1, IREST=0,
TIMLIM=999999, IG=33447, IBELLY=0, NTR=0,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SC EE=1.2,
NSTLIM=1, DT=0.001, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=0.001, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=1, NTWX=1,
NMROPT=0,
&end
&ewald,
a=62.6631520, b=51.4624957, c=51.0325726,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum_tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end

```

ei300K8

```

DNA
&cntrl
IMIN=0, NTX=7, IREST=1,
TIMLIM=999999, IG=77334, IBELLY=0, NTR=0,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SC EE=1.2,
NSTLIM=5000, DT=0.002, NDFMIN=6, NTCM=1, NSCM=100,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,

```



```

NTPR=50, NTWX=500,
NMROPT=0,
&end
&ewald,
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54,order=4, dsum_tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end

```

prodstb

```

DNA
&cntrl
IMIN=0, NTX=7, IREST=1,
TIMLIM=999999, IG=99775, IBELLY=0, NTR=0,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=5000, DT=0.002, NDFMIN=6, NTCM=1, NSCM=100,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=0,
&end
&ewald
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54,order=4, dsum_tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end

```

ri300K_Maxwell2

```

DNA
&cntrl
IMIN=0, NTX=1, IREST=0,
TIMLIM=999999, IG=79681, IBELLY=0, NTR=0,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=1, DT=0.001, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=0.001, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=1, NTWX=1,
NMROPT=0,

```

```

&end
&ewald,
a=62.6975842, b=51.4907733, c=51.0606141,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum__tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end

```

prodstb

```

DNA
&cntrl
IMIN=0, NTX=7, IREST=1,
TIMLIM=999999, IG=99775, IBELLY=0, NTR=0,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=5000, DT=0.002, NDFMIN=6, NTCM=1, NSCM=100,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=0,
&end
&ewald
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum__tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end

```

ri300K_Maxwell3

```

DNA
&cntrl
IMIN=0, NTX=1, IREST=0,
TIMLIM=999999, IG=28272, IBELLY=0, NTR=0,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SCEE=1.2,
NSTLIM=1, DT=0.001, NDFMIN=0, NTCM=0, NSCM=0,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=0.001, TAUTP=0.2,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=1, NTWX=1,
NMROPT=0,
&end
&ewald,

```

```

a=62.5470899, b=51.3671788, c=50.9380522,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum_tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end

```

dn300K

```

DNA
&cntrl
NTX=7, IREST=1,
IG=99775, IBELLY=0, NTR=0,
NTB=2,
DIELC=1, CUT=9.0, SCNB=2.0, SC EE=1.2,
NSTLIM=250000, DT=0.002, NDFMIN=6, NTCM=1, NSCM=100,
HEAT=0.0, TEMPI=300.0, TEMP0=300.0, NTT=1, DTEMP=2.0, TAUTP=0.2,
VLIMIT=20.0,
NTP=1, PRES0=1.0, COMP=44.6, TAUP=0.2,
NTC=2, TOL=0.0005, NTF=2,
NTPR=50, NTWX=500,
NMROPT=0,
&end
&ewald
a=65.2688, b=53.6024, c=53.1546,
nfft1=64, nfft2=54, nfft3=54, order=4, dsum_tol=0.000001, &end
&wt TYPE='END', &end
&rst iat=0, &end

```

Analyse.pl

```

#!/usr/bin/perl -w
# Programme servant a mettre en forme des donnees
# afin de les visualiser facilement avec XmGrace
# version de developpement 28052003
# Special thanks to Paride Papadia for debugging info
# Special thanks to Markus Drumm for debugging info
use strict ;
# Tout d'abord les repertoires de travail
my $carnalout="./" ;
my $c5selout="../c5_sel_out/" ;
my $repsortie="Illustrations/" ;
my $suffixeglobal="_moy" ;
# Ensuite la sequence du dinucleotide ainsi que le titre des graphes
# TOUJOURS dans le sens 5'->3'
my @brin1=('G','C','C','G*','G*','G','T','C','G','C') ;

```

```

my @brin2=('G','C','G','A','C','C','C','G','G','C');
my $titreprincipal="GGG-Pt";
my $soustitreprincipal="15 ns en eau explicite - parm99";
# Quels parametres vont etre representes?.
my @squelette=("alpha","beta","gamma","delta","epsilon","zeta","epmoinsze","Phase","Amplitude","chi","Incl");
my @interbase=("Twist_Omega","Roll_rho","Tilt_tau","Shift_Dx","Slide_Dy","Rise_Dz");
my @intrapairebase1=("Opening_sigma","Buckle_kappa","Stretch_Sy");
my @intrapairebase2=("Propel_omega","Shear_Sx","Stagger_Sz");
my @divers1=("Atip");
my @divers2=("Angle");
my @ungraphe=("Kink");
# Pour avoir une echelle automatique sur les axes :
# soit aucune : mettre la valeur "onread none", les valeurs définies plus bas seront utilisées
# soit pour x : mettre la valeur "xaxes"
# soit pour y : mettre la valeur "yaxes"
my $autoscale="yaxes";
# Pour ne représenter qu'une fraction du dinucléotide :
# Il faut au moins un décalage de un pour les pas inter paire de base
my $brin1debut=1;
my $brin1fin=10;
# Pour mettre de la couleur dans les représentations.
my $couleurtitre=4;
my $couleursoustitre=2;
# Format de sortie souhaité : png, ps.
my $format="ps";
#
# A partir d'ici il n'y a plus rien à changer ....
#
my $pwd=$ENV{PWD};
my $nbgraph=@brin1;
# Pour faire un peu le ménage
if (! -d $repsortie )
{
mkdir $repsortie;
}
my %min=("alpha"=>"-140",
"beta"=>"90",
"gamma"=>"0",
"delta"=>"60",
"epsilon"=>"120",
"zeta"=>"60",
"epmoinsze"=>"-130",
"Phase"=>"-60",
"Amplitude"=>"20",
"chi"=>"150",
"Atip"=>"-20",

```

```

"Angle"=>"0",
"Epsilon_amplitude"=>"150",
"Inclin_eta"=>"-30",
"Kink"=>"0",
"Propel_omega"=>"-50",
"Opening_sigma"=>"-20",
"Rise_Dz"=>"2",
"Roll_rho"=>"-50",
"Shear_Sx"=>"-2",
"Stretch_Sy"=>"-1",
"Shift_Dx"=>"-3",
"Slide_Dy"=>"-3",
"Stagger_Sz"=>"-2",
"Buckle_kappa"=>"-45",
"Tilt_tau"=>"-40",
"Tip_theta"=>"-40",
"Twist_Omega"=>"10",
"Temps"=>"0"
);
my %max=("alpha"=>"0",
"beta"=>"270",
"gamma"=>"320",
"delta"=>"180",
"epsilon"=>"360",
"zeta"=>"330",
"epmoinsze"=>"230",
"Phase"=>"240",
"Amplitude"=>"60",
"chi"=>"360",
"Atip"=>"50",
"Angle"=>"30",
"Epsilon_amplitude"=>"300",
"Inclin_eta"=>"60",
"Kink"=>"100",
"Propel_omega"=>"30",
"Opening_sigma"=>"30",
"Rise_Dz"=>"5",
"Roll_rho"=>"50",
"Shear_Sx"=>"2",
"Stretch_Sy"=>"2",
"Shift_Dx"=>"3",
"Slide_Dy"=>"3",
"Stagger_Sz"=>"2",
"Buckle_kappa"=>"60",
"Tilt_tau"=>"40",
"Tip_theta"=>"20",

```

```

"Twist_Omega"=>"50",
"Temps"=>"15000"
);
my %typesortie=("ps"=>"Postscript", "png"=>"PNG", "pdf"=>"PDF");
sub entete
{
my (@t)=@_ ;
while (@t)
{
my ($fichier)=shift @t ;
my $sortie=$repsortie.$fichier.".batch" ;
open(FIC,"> $sortie") or die ("Ouverture du fichier : $!\n");
print FIC "# Creation de fichiers pour XmGrace
page layout portrait
page size 595, 842
page scroll 5%
page inout 5%
background color 0\n";
close FIC ;
}
}
sub nombregraphes
{
# Création de la matrice des graphes
my ($nb,@t)=@_ ;
while (@t)
{
my ($fichier)=shift @t ;
my $sortie=$repsortie.$fichier.".batch" ;
open(FIC,">> $sortie") or die ("Ouverture du fichier :$!\n");
print FIC "arrange ($nb,2,0.15,.25,.2,ON,OFF,OFF,ON)\n";
my $i=0;
while ( $i <= $nb*2 )
{
print FIC "G$i HIDDEN ON\n";
$i+=1 ;
}
close FIC ;
}
}
sub titregraphe
{
my ($nb,@t)=@_ ;
while (@t)
{
my ($fichier)=shift @t ;

```

```

my $sortie=$repsortie.$fichier.".batch";
open(FIC,">> $sortie") or die ("Ouverture du fichier :$!\n");
print FIC "#Pour le graphe de titre
g$nb on
with g$nb
view xmin 0.05
view xmax 0.95
view ymin 0.05
view ymax 1.30
title \"$titreprincipal\"
title font 0
title size 1.500000
title color $couleurtitre
subtitle \"$soustitreprincipal - $fichier -\"
subtitle font 0
subtitle size 1.000000
subtitle color $couleursoustitre
xaxis off
xaxis bar off
yaxis off
yaxis bar off
yaxis tick off
yaxis ticklabel off
legend off
frame type 0
frame linestyle 0
frame color 0\n\n";
close FIC;
}
}
sub titrelocalgraphe
{
my ($nb,@t)=@_ ;
while (@t)
{
my ($fichier)=shift @t;
my $sortie=$repsortie.$fichier.".batch";
open(FIC,">> $sortie") or die ("Ouverture du fichier :$!\n");
print FIC "with g$nb
subtitle \"$fichier\" \n subtitle font 0\n subtitle size 1.00\n\n";
close FIC;
}
}
sub graphe
{
my ($nographe, $suffixe, $nomx, $affichx, $nomy, $posy, $colonne, @t)=@_ ;

```

```

while (@t)
{
my ($fichier)=shift @t ;
my $sortie=$repsortie.$fichier.".batch";
open(FIC,">> $sortie") or die ("Ouverture du fichier :$!\n");
print FIC "# Pour le graphe $nographe\n";
print FIC "g$nographe on\n";
print FIC "with g$nographe\n";
print FIC "read block \"$c5selout$fichier$suffixe$suffixeglobal\"\n";
print FIC "block xy \"1 :$colonne\"\n";
#Coordonnees globales
print FIC "world xmin $min{Temps}\n";
print FIC "world xmax $max{Temps}\n";
print FIC "world ymin $min{$fichier}\n";
print FIC "world ymax $max{$fichier}\n";
print FIC "autoscale $autoscale\n";
# Caracteristiques de l'axe x
print FIC "xaxis on\n";
print FIC "xaxis label \"$nomx\"\n";
print FIC "xaxis type zero false\n";
print FIC "xaxis tick on\n";
print FIC "xaxis tick default 6\n";
print FIC "xaxis tick place rounded true\n";
print FIC "xaxis tick major size 0.25\n";
print FIC "xaxis ticklabel $affichx \n";
print FIC "xaxis ticklabel char size 0.5\n";
# Caracteristiques de l'axe y
print FIC "yaxis on\n";
print FIC "yaxis label \"$nomy\"\n";
print FIC "yaxis label place $posy\n";
print FIC "yaxis label layout perp\n";
print FIC "yaxis tick on\n";
print FIC "yaxis tick place rounded true\n";
print FIC "yaxis tick major size 0.25\n";
print FIC "yaxis ticklabel on\n";
print FIC "yaxis ticklabel char size 0.5\n\n";
close FIC ;
}
}
sub sortie
{
my (@t)=@_ ;
while (@t)
{
my ($fichier)=shift @t ;
my $sortie=$repsortie.$fichier.".batch";

```



```

open(FIC,"> > $sortie") or die ("Ouverture du fichier :$!\n");
print FIC "#Pour la sortie en image\n";
print FIC "hardcopy device \"${typesortie}{format}\" \"\n\";
print FIC "print to \"${fichier.format}\" \"\n\";
print FIC "print\n";
close FIC;
}
}
sub impression
{
my (@t)=@_ ;
while (@t)
{
my ($fichier)=shift @t;
my $sortie=$repsortie.$fichier.".batch";
system "gracebat -nosafe -batch $fichier.batch";
}
}
# Pour imprimer les paramètres du squelette
my $finnuc=$brin1fin-$brin1debut+1;
entete (@squelette);
nombregraphes ($finnuc,@squelette);
titregraphe ($finnuc*2,@squelette);
for ( my $graph=0;$graph<$finnuc-1;$graph++ )
{
my $nuc=$graph+$brin1debut;
my $nomy=$brin1[$nuc-1].'\s'.$nuc;
graphe ($graph, '_s1_5p3p.out', ", 'off', $nomy, 'normal', ($nuc+1), @squelette);
}
my $nomy=$brin1[$finnuc+$brin1debut-2].'\s'($finnuc+$brin1debut-1);
graphe ($finnuc-1, '_s1_5p3p.out', 'Temps (ps)', 'on', $nomy, 'normal', ($finnuc+$brin1debut),
@squelette);
my $intervalle=($nbgraph-$brin1fin);
$nomy=$brin2[$intervalle].'\s'($intervalle+$nbgraph+1);
graphe ($finnuc, '_s2_5p3p.out', 'Temps (ps)', 'on', $nomy, 'opposite', ($intervalle+2), @sque-
lette);
my $intervallefin=2*$finnuc;
my $nuc=$intervalle+1;
my $graph=$finnuc+1;
while ( $graph<$intervallefin)
{
my $nomy=$brin2[$nuc].'\s'($nbgraph+$nuc+1);
graphe ($graph, '_s2_5p3p.out', ", 'off', $nomy, 'opposite', ($nuc+2), @squelette);
$nuc+=1;
$graph+=1;
}

```

```

sortie (@squelette);
# Pour imprimer les paramètres Inter paires de bases
$finnuc=$brin1fin-$brin1debut;
entete (@interbase);
nombregraphes ($finnuc,@interbase);
titregraphe ($finnuc*2,@interbase);
for ( my $graph=0;$graph<$finnuc-1;$graph++)
{
my $nuc=$graph+$brin1debut;
my $nomy=$brin1[$nuc-1].'\s'.$nuc.'\\N'.'/'. $brin1[$nuc].'\s'($nuc+1);
graphe ($graph, '_s1_5p3p.out', ", 'off', $nomy, 'normal', $nuc+1, @interbase);
}
$nomy=$brin1[$finnuc+$brin1debut-2].'\s'($finnuc+$brin1debut-1).'\N'.'/'. $brin1[$finnuc+$brin1debut-
1].'\s'($finnuc+$brin1debut);
graphe ($finnuc-1, '_s1_5p3p.out', 'Temps (ps)', 'on', $nomy, 'normal', $finnuc+$brin1debut,
@interbase);
$intervalle=($nbgraph-$brin1fin);
$nomy=$brin2[$intervalle].'\s'($intervalle+$nbgraph+1).'\N'.'/'. $brin2[$intervalle+1].'\s'($intervalle+$nbg
raphe ($finnuc, '_s2_5p3p.out', 'Temps (ps)', 'on', $nomy, 'opposite', $intervalle+2, @interbase);
$intervallefin=2*$finnuc;
$nuc=$intervalle+1;
$graph=$finnuc+1;
while ( $graph<$intervallefin)
{
my $nomy=$brin2[$nuc].'\s'($nbgraph+$nuc+1).'\N'.'/'. $brin2[$nuc+1].'\s'($nbgraph+$nuc+2);
graphe ($graph, '_s2_5p3p.out', ", 'off', $nomy, 'opposite', $nuc+2, @interbase);
$nuc+=1;
$graph+=1;
}
sortie (@interbase);
# Pour imprimer les paramètres Intra paires de bases
$finnuc=$brin1fin-$brin1debut+1;
entete (@intrapairebase1);
nombregraphes ($finnuc,@intrapairebase1);
titregraphe ($finnuc*2,@intrapairebase1);
for ( my $graph=0;$graph<$finnuc-1;$graph++)
{
my $nuc=$graph+$brin1debut;
my $nomy=$brin1[$nuc-1].'\s'.$nuc.'\\N :'. $brin2[$nbgraph-$nuc].'\s'($nbgraph*2+1-$nuc);
graphe ($graph, '.out', ", 'off', $nomy, 'normal', ($nuc+1), @intrapairebase1);
}
titrelcalgraphe (0,@intrapairebase1);
$nomy=$brin1[$finnuc+$brin1debut-2].'\s'($finnuc+$brin1debut-1).'\N :'. $brin2[$nbgraph-$brin1fin].'\s'($n
$brin1fin+1);
graphe ($finnuc-1, '.out', 'Temps (ps)', 'on', $nomy, 'normal', ($finnuc+$brin1debut), @intrapai-
rebase1);

```

```

$intervalle=($nbgraph-$brin1fin);
$nomy=$brin1[$nbgraph-$intervalle-1].'\s'.($nbgraph-$intervalle).'\N:'. $brin2[$intervalle].'\s'.($intervalle+$nbgr
graphe ($finnuc, 'out', 'Temps (ps)', 'on', $nomy, 'opposite', ($nbgraph-$intervalle+1), @intra-
pairebase2);
$intervallefin=2*$finnuc;
$nuc=$intervalle+1;
$graph=$finnuc+1;
while ( $graph<$intervallefin)
{
my $nomy=$brin1[$nbgraph-$nuc-1].'\s'.($nbgraph-$nuc).'\N:'. $brin2[$nuc].'\s'.($nbgraph+$nuc+1);
graphe ($graph, 'out', ", 'off', $nomy, 'opposite', ($nbgraph-$nuc+1), @intrapairebase2);
$nuc+=1;
$graph+=1;
}
titrelocalgraphe ($graph-1,@intrapairebase2);
sortie (@intrapairebase1);
# Pour imprimer les paramètres Inter paires de bases
$finnuc=$brin1fin-$brin1debut;
entete (@interbase);
nombregraphes ($finnuc,@interbase);
titregraphe ($finnuc*2,@interbase);
for ( my $graph=0;$graph<$finnuc-1;$graph++ )
{
my $nuc=$graph+$brin1debut;
my $nomy=$brin1[$nuc-1].'\s'. $nuc.'\\N'.'\/'. $brin1[$nuc].'\s'.($nuc+1);
graphe ($graph, '_s1_5p3p.out', ", 'off', $nomy, 'normal', $nuc+1, @interbase);
}
$nomy=$brin1[$finnuc+$brin1debut-2].'\s'.($finnuc+$brin1debut-1).'\N'.'\/'. $brin1[$finnuc+$brin1debut-
1].'\s'.($finnuc+$brin1debut);
graphe ($finnuc-1, '_s1_5p3p.out', 'Temps (ps)', 'on', $nomy, 'normal', $finnuc+$brin1debut,
@interbase);
$intervalle=($nbgraph-$brin1fin);
$nomy=$brin2[$intervalle].'\s'.($intervalle+$nbgraph+1).'\N'.'\/'. $brin2[$intervalle+1].'\s'.($intervalle+$nbgrap
graphe ($finnuc, '_s2_5p3p.out', 'Temps (ps)', 'on', $nomy, 'opposite', $intervalle+2, @interbase);
$intervallefin=2*$finnuc;
$nuc=$intervalle+1;
$graph=$finnuc+1;
while ( $graph<$intervallefin)
{
my $nomy=$brin2[$nuc].'\s'.($nbgraph+$nuc+1).'\N'.'\/'. $brin2[$nuc+1].'\s'.($nbgraph+$nuc+2);
graphe ($graph, '_s2_5p3p.out', ", 'off', $nomy, 'opposite', $nuc+2, @interbase);
$nuc+=1;
$graph+=1;
}
sortie (@interbase);
# Pour imprimer les paramètres Divers

```

```

$finnuc=$brin1fin-$brin1debut+1;
entete (@divers1);
nombregraphes ($finnuc-1,@divers1);
titregraphe ($finnuc*2-2,@divers1);
for ( my $graph=0;$graph<$finnuc-2;$graph++ )
{
my $nuc=$graph+$brin1debut;
my $nomy=$brin1[$nuc-1].'\s'.$nuc.'\N /'.$brin1[$nuc].'\s'($nuc+1);
graphe ($graph, 'out', ", 'off', $nomy, 'normal', ($nuc+1), @divers1);
}
titrelocalgraphe (0,@divers1);
$nomy=$brin1[$finnuc+$brin1debut-3].'\s'($finnuc+$brin1debut-2).\N /'.$brin1[$finnuc+$brin1debut-
2].'\s'($finnuc+$brin1debut-1);
graphe ($finnuc-2, 'out', 'Temps (ps)', 'on', $nomy, 'normal', ($finnuc+$brin1debut-1), @divers1);
$intervalle=($nbgraph-$brin1fin);
# $nomy=$brin1[$intervalle+1].'\s'($nbgraph-$intervalle-1).\N /'.$brin1[$intervalle+2].'\s'($nbgraph-
$intervalle);
$nomy=$brin1[$nbgraph-$intervalle-2].'\s'($nbgraph-$intervalle-1).\N /'.$brin1[$nbgraph-$intervalle-
1].'\s'($nbgraph-$intervalle);
graphe ($finnuc-1, 'out', 'Temps (ps)', 'on', $nomy, 'opposite', ($intervalle+2), @divers2);
$intervallefin=2*$finnuc-2;
$nuc=$intervalle+1;
$graph=$finnuc;
while ( $graph<$intervallefin)
{
my $nomy=$brin1[$nbgraph-$nuc-2].'\s'($nbgraph-$nuc-1).\N /'.$brin1[$nbgraph-$nuc-1].'\s'($nbgraph-
$nuc);
graphe ($graph, 'out', ", 'off', $nomy, 'opposite', ($nuc+2), @divers2);
$nuc+=1;
$graph+=1;
}
titrelocalgraphe ($graph-1,@divers2);
sortie (@divers1);
# Pour imprimer un graphe
entete (@ungraphe);
nombregraphes (5,@ungraphe);
titregraphe (9,@ungraphe);
graphe (2, 'out', 'Temps (ps)', 'on', "Kink", 'normal', 2, @ungraphe);
sortie (@ungraphe);
#Pour l'impression de tous les paramètres
chdir ($repsortie);
my $tour=@intrapairebase2;
my $debut=0;
while ($debut <$tour)
{
my $fich1=$intrapairebase1[$debut].'.batch';

```

```

my $fich2=$intrapairebase2[$debut].'.batch';
system 'cat $fich1 $fich2 > toto.batch';
system 'mv toto.batch $fich1';
$debut+=1;
}
$debut=0;
$tour=@divers2;
while ($debut < $tour)
{
my $fich1=$divers1[$debut].'.batch';
my $fich2=$divers2[$debut].'.batch';
system 'cat $fich1 $fich2 > toto.batch';
system 'mv toto.batch $fich1';
$debut+=1;
}
impression (@squelette);
impression (@interbase);
impression (@intrapairebase1);
impression (@divers1);
impression (@ungraphe);
while ( @intrapairebase2 )
{
my ($fichier)=shift @intrapairebase2;
my $sortie=$fichier.'.batch';
my $ret=unlink $sortie;
}
while ( @divers2 )
{
my ($fichier)=shift @divers2;
my $sortie=$fichier.'.batch';
my $ret=unlink $sortie;
}
chdir ($pwd);

```